



HHS Public Access

Author manuscript

Biom J. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

Biom J. 2015 September ; 57(5): 766–776. doi:10.1002/bimj.201300001.

Log-gamma linear-mixed effects models for multiple outcomes with application to a longitudinal glaucoma study

Peng Zhang^{*,1}, Dandan Luo², Pengfei Li³, Lucie Sharpsten⁴, and Felipe A. Medeiros⁴

¹Department of Mathematics, Zhejiang University, 86 Zheda Road, Hangzhou, Zhejiang 310012, China

²Bank of Montreal, Toronto M5X 1A1, Canada

³Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

⁴Hamilton Glaucoma Center, Department of Ophthalmology, University of California, San Diego, La Jolla, CA 92093, USA

Abstract

Glaucoma is a progressive disease due to damage in the optic nerve with associated functional losses. Although the relationship between structural and functional progression in glaucoma is well established, there is disagreement on how this association evolves over time. In addressing this issue, we propose a new class of non-Gaussian linear-mixed models to estimate the correlations among subject-specific effects in multivariate longitudinal studies with a skewed distribution of random effects, to be used in a study of glaucoma. This class provides an efficient estimation of subject-specific effects by modeling the skewed random effects through the log-gamma distribution. It also provides more reliable estimates of the correlations between the random effects. To validate the log-gamma assumption against the usual normality assumption of the random effects, we propose a lack-of-fit test using the profile likelihood function of the shape parameter. We apply this method to data from a prospective observation study, the Diagnostic Innovations in Glaucoma Study, to present a statistically significant association between structural and functional change rates that leads to a better understanding of the progression of glaucoma over time.

Keywords

Functional progression; Glaucoma; Log-gamma distribution; Markov chain Monte Carlo; Multivariate longitudinal data; Profile likelihood; Structural progression

*Corresponding author: pengz@zju.edu.cn, Phone: +86-571-87953867, Fax: +86-571-87953794.

Additional supporting information may be found in the online version of this article at the publisher's web-site

Conflict of interest

The authors have declared no conflict of interest.

1 Introduction

Glaucoma is an eye disease characterized by progressive neuroretinal rim thinning, excavation, and loss of the retinal nerve fiber layer. These structural changes are usually accompanied by functional losses. Although there is a relationship between structural and functional damage in glaucoma, the precise association and the evolution of this association over time are still unclear. For example, it is not yet known whether the glaucomatous structure and function develop at about the same rate, and how rates of change in structure are related to those of function at different stages of the disease. The evolution of association in rates of change is essential to enhance the understanding of the glaucomatous process and to determine the relative utility of tests for structure and function in monitoring different stages of glaucoma. Many studies have investigated the structural and functional relationship in glaucoma (Harwerth et al., 2010; Malik et al., 2012; Medeiros et al., 2012). These studies have shown different degrees of association between quantitative structural measures derived from imaging technologies and psychophysical tests such as standard automated perimetry (SAP). Similarly, large disagreements have been reported in longitudinal studies evaluating glaucoma progression (Keltner et al., 2006; Strouthidis et al., 2006).

The association between the functional and structural changes in glaucoma can be modeled by studying the relationships among the subject-specific effects in mixed effects models for the longitudinal data. Statistically, we study the correlation of the random intercepts between the two change responses, and the correlation of the random slopes of time effects. Multivariate models are used when responses for two or more response variables are observed over time for each individual in a longitudinal study; such data are commonly collected in the health sciences and epidemiological studies. There is much literature on multivariate longitudinal data. For example, Reinsel (1984) considered the random effects model for multiple outcomes with a complete and balanced design. Shah et al. (1997) extended linear-mixed models to allow for multiple longitudinal outcomes in the case where the number and timing of observations may differ from individual to individual. Roy and Lin (2000) extended latent variable models to multivariate longitudinal data. There are two sources of within-subject correlation that should be reflected in the multiple-outcomes model: (i) among different outcomes; (ii) among repeated measures of the same outcome over time. Although one can perform separate analyses of the two outcomes, this does not address the main question of interest: how overall treatment practices have changed over time. Moreover, the analysis will have increased power if information from all of the outcomes is used (Pocock et al., 1987).

Our research is motivated by analyzing data obtained from a prospective observational study, the Diagnostic Innovations in Glaucoma Study (DIGS). The DIGS started in Sep. 2005 and finished recruitment in Aug. 2011. It was carried out at the Hamilton Glaucoma Center of the Department of Ophthalmology, University of California, San Diego. The DIGS was designed to investigate the longitudinal structure and function relationship in glaucoma. Two responses, the visual field index (VFI) and the global average thickness of the retinal nerve fiber layer (RNFL), were collected for 203 patients at 6-month intervals.

The first response characterizes the functional loss, while the second characterizes the structural changes. Section 5 gives more information about the study and the responses.

In a preliminary analysis (see Supporting Information Fig. S1), we observed that the intercepts from a within-subject linear regression analysis fitting the logit-transformed VFI to time had a left-skewed distribution, while the within-subject distributions of the slopes for VFI and those of the intercepts and slopes for RNFL were roughly symmetric. To accommodate the skewness of one of the random effects and the associations among the multivariate responses, we propose extending the mixed effects model for longitudinal data of Zhang et al. (2008) to a bivariate linear-mixed effects model where the random effects are assumed to follow the log-gamma distribution. For the multivariate response model, the usual distributional assumption for the random effects is multivariate normal. However, there are only a few results on the consequences for statistical inference of misspecifying the random-effects distribution. This may impact the bias and variance of the parameter estimation and the power of hypothesis-testing procedures. Ghosh et al. (2007) developed a Bayesian approach to the bivariate-mixed effects model via a multivariate skew-normal distribution. However, there is no closed form of the marginal distribution of the responses when the random effects are assumed to be nonnormal. In this paper, we present a nonnormal linear-mixed effects model for multiple outcomes, which is feasible mathematically and computationally less complex. It is applied to the bivariate longitudinal study of glaucoma to investigate the association between the structural and functional progressions in the patients' eyes.

The remainder of this article is organized as follows. Section 2 presents a nonnormal multivariate linear-mixed effects model. Section 3 discusses the inferences of the fixed effects, variance components, and random effects. We use a Gauss–Newton algorithm to calculate the maximum likelihood estimates (MLEs) of the parameters and the Markov chain Monte Carlo (MCMC) method for the inference of the random effects. Section 4 presents a lack-of-fit test, based on the fact that the limiting distribution of the log-gamma distribution is normal, to test the adequacy of the log-gamma assumption. Section 5 presents the tests and the results, and Section 6 provides a discussion. Two figures and a brief introduction to the log-gamma distribution are given in the Supplementary Material.

2 Model and assumptions

In this section, we define a general multivariate linear-mixed effects model assuming that the skewed random effect follows a log-gamma distribution. Consider a multivariate longitudinal study with p responses and N subjects. For $i = 1, \dots, N$ and $k = 1, \dots, p$, let

$\mathbf{Y}_{ik} = (Y_{i1k}, \dots, Y_{in_{ik}k})'$ denote the collection of n_{ik} observations for the k -th outcome of the i -th subject. We assume a linear-mixed effects model for \mathbf{Y}_{ik} of the form

$$\mathbf{Y}_{ik} = X_{ik}\boldsymbol{\beta}_k + Z_{ik}\mathbf{b}_{ik} + \boldsymbol{\varepsilon}_{ik}, i=1, \dots, N; k=1, \dots, p. \quad (1)$$

Here $\boldsymbol{\beta}_k$ is a vector of fixed effects for the k -th response, X_{ik} is the design matrix of fixed effects $\boldsymbol{\beta}_k$, \mathbf{b}_{ik} is a vector of random effects for the k -th response, and Z_{ik} is the design matrix of random effects \mathbf{b}_{ik} .

Further, let $\mathbf{Y}_i = (\mathbf{Y}'_{i1}, \dots, \mathbf{Y}'_{ip})'$ be a stacked vector of p responses for the i -th subject, $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \dots, \boldsymbol{\beta}'_p)'$ be a stacked vector of all fixed effects, $\mathbf{b}_i = (\mathbf{b}'_{i1}, \dots, \mathbf{b}'_{ip})'$ be a stacked vector of all random effects, $\mathbf{X}_i = \text{diag}(\mathbf{X}_{i1}, \dots, \mathbf{X}_{ip})$ be a block diagonal design matrix for all fixed effects, $\mathbf{Z}_i = \text{diag}(\mathbf{Z}_{i1}, \dots, \mathbf{Z}_{ip})$ be a block diagonal design matrix for all random effects, and $\boldsymbol{\varepsilon}_i = (\boldsymbol{\varepsilon}'_{i1}, \dots, \boldsymbol{\varepsilon}'_{ip})'$ be a stacked vector of error terms. Then the linear-mixed effects model in (1) can be written

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \boldsymbol{\varepsilon}_i, i = 1, \dots, N. \quad (2)$$

We assume that \mathbf{Y}_i is independent of \mathbf{Y}_j for $i \neq j$. That is, the stacked response vectors of two different subjects are independent. We also assume that \mathbf{Y}_{ik} and \mathbf{Y}_{il} , $k \neq l$ are conditionally independent given the random effects \mathbf{b}_i . That is, for the same subject, the k -th and l -th response vectors are conditionally independent, given the random effects for that subject. Further, we assume that $E(\mathbf{b}_i) = \mathbf{0}$ and $E(\boldsymbol{\varepsilon}_i) = \mathbf{0}$.

To model the skewness of the random effects, we rearrange the components in \mathbf{b}_i and divide them into two parts. Denote the reordered vector $\mathbf{b}_i^* = (b_{i1}^*, \dots, b_{iq}^*)'$. The first q_0 components have skewed distributions, and the remaining $q - q_0$ components are symmetric. The reordering also results in the interchange of the columns in the design matrix \mathbf{Z}_i according to the subscripts of \mathbf{b}_i^* . To allow correlated random intercepts and random slopes, we define $\mathbf{b}_i^* = \mathbf{R} \mathbf{s}_i^*$, where $\mathbf{s}_i^* = (s_{i1}^*, \dots, s_{iq}^*)'$, and the linear transformation matrix \mathbf{R} is an upper triangular matrix with unknown entries r_{ij} , $1 \leq i < j \leq q$, and 1's on the diagonal. The components of the vector \mathbf{s}_i^* are independent of each other. Expressing the transformation in matrix notation gives

$$\begin{pmatrix} b_{i1}^* \\ \vdots \\ b_{iq_0}^* \\ \vdots \\ b_{iq}^* \end{pmatrix} = \begin{pmatrix} 1 & \dots & r_{1q_0} & \dots & r_{1q} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & r_{q_0q} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} s_{i1}^* \\ \vdots \\ s_{iq_0}^* \\ \vdots \\ s_{iq}^* \end{pmatrix}. \quad (3)$$

In (3), we specify that just one random effect in \mathbf{s}_i^* , $s_{iq_0}^*$, is from the log-gamma distribution. See Section 2 of the Supplementary Material for a brief introduction to the log-gamma distribution. From (3), we observe that $b_{i1}^*, \dots, b_{iq_0}^*$ are linear combinations of $s_{iq_0}^*$ and other random effects in \mathbf{s}_i^* . By specifying the distribution of $s_{iq_0}^*$ to be skewed, we can model $b_{i1}^*, \dots, b_{iq_0}^*$ as skewed random effects. Allowing multiple random effects in \mathbf{s}_i^* to be from the log-gamma distribution would complicate the calculation of the MLEs of the unknown parameters. Moreover, in our real-data analysis, there is only one skewed random effect, that is the random intercept for the logit-transformed VFI. In (3), we just need to set $q_0 = 1$. More details can be found in Section 5.

The mean of the log-gamma random variable $s_{iq_0}^*$ is set to zero, and the remaining components of \mathbf{s}_i^* are assumed to have a multivariate normal distribution with mean $\mathbf{0}$ and diagonal covariance matrix \mathbf{G} . Thus, from (3), $E(\mathbf{b}_i^*)=0$ and so $E(\mathbf{b}_i) = \mathbf{0}$. If the distribution of the random effect $b_{iq_0}^*$ is positively skewed, then a minus log-gamma distribution should be selected for $s_{iq_0}^*$ because the log-gamma distribution is negatively skewed (see Supporting Information Fig. S3). For the other skewed components, the sign of the coefficient of s_{ij}^* , $1 \leq j < q_0$, determines the direction of skewness. The advantage of introducing the linear transformation matrix \mathbf{R} rather than specifying the correlated random effects directly is that this gives a relatively simple model and also makes it easy to implement the log-gamma mixed-effects model.

3 Inference

In this section, we discuss how to calculate the estimates of the unknown parameters and the random effects. We denote the unknown parameters as $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mathbf{G}, \mathbf{R}, \kappa, \Sigma_i)'$, where $\boldsymbol{\beta}$ contains the regression coefficients, \mathbf{G} contains the covariance components, \mathbf{R} contains the parameters of the linear transformation matrix in (3), κ is the log-gamma shape parameter, and $\Sigma_i = \text{Cov}(\boldsymbol{\varepsilon}_i) = \text{diag}(\sigma_1^2 I_{n_{i_1}}, \sigma_2^2 I_{n_{i_2}}, \dots, \sigma_p^2 I_{n_{i_p}})$ contains the conditional covariance parameters. The estimation and inference of $\boldsymbol{\theta}$ are based on the marginal likelihood function, which is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^N \int \{ \int \cdots \int \phi(\mathbf{y}_i | \mathbf{s}_i^*) \cdot \phi(s_{i_1}^*, \dots, s_{iq_0-1}^*, s_{iq_0+1}^*, \dots, s_{i_q}^*) ds_{i_1}^* \cdots ds_{iq_0-1}^* ds_{iq_0+1}^* \cdots ds_{i_q}^* \} \cdot f(s_{iq_0}^*) ds_{iq_0}^*, \quad (4)$$

where $f(s_{iq_0}^*)$ represents the density function of the standard log-gamma distribution (see Eq. (1) of the Supplementary Material). The MLE of $\boldsymbol{\theta}$ is defined to be $\hat{\boldsymbol{\theta}} = \arg \sup_{\boldsymbol{\theta}} L(\boldsymbol{\theta})$. By verifying the regularity conditions in Weiss (1973) and Jiang (2001), we can show that the MLE $\hat{\boldsymbol{\theta}}$ is consistent and asymptotically normal, with the asymptotic covariance matrix equal to the inverse of the Fisher information matrix. Because there is no closed-form expression for the integration (4), we need numerical integration techniques to maximize the likelihood function.

To find $\hat{\boldsymbol{\theta}}$ numerically, we use a Gauss–Newton algorithm (Ruppert, 2005), in which the Fisher information matrix is approximated by $\mathbf{B}(\boldsymbol{\theta}) = \sum_{i=1}^N i_i(\mathbf{y}_i; \boldsymbol{\theta}) i_i(\mathbf{y}_i; \boldsymbol{\theta})'$, where $i_i(\mathbf{y}_i; \boldsymbol{\theta})$ is the log-likelihood contributed by the i -th subject and $i_i(\mathbf{y}_i; \boldsymbol{\theta})$ is the first derivative of $l_i(\mathbf{y}_i; \boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. Thus, only the first-order derivatives of the log likelihood are involved. The key step of this algorithm is to halve the value of δ , which guarantees a steady increase in the likelihood between iterations. To be more precise, the $(k + 1)$ -th iteration calculates

$$\boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^k + \delta \{ \mathbf{B}(\boldsymbol{\theta}^k) \}^{-1} i(\boldsymbol{\theta}^k),$$

where δ is the step-halving term and $i(\boldsymbol{\theta}) = \sum_{i=1}^N i_i(\mathbf{y}_i; \boldsymbol{\theta})$. The algorithm starts with $\delta = 1$ and halves δ until the log-likelihood function $l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$ satisfies $l(\boldsymbol{\theta}^{k+1}) > l(\boldsymbol{\theta}^k)$. It stops when an increase in the likelihood is no longer possible, or the difference between two consecutive updates is smaller than a prespecified precision. The starting values are obtained by fitting a bivariate linear-mixed model assuming normal random effects using the SAS procedure MIXED, as discussed by Thiébaud et al. (2002).

Once $\hat{\boldsymbol{\theta}}$ has been obtained, we use it to estimate the random effects. The posterior distribution of \mathbf{b}_i^* is given by

$$f(\mathbf{b}_i^* | \mathbf{y}_i) = \frac{\phi(\mathbf{y}_i | \mathbf{b}_i^*) f(\mathbf{b}_i^*)}{\int \phi(\mathbf{y}_i | \mathbf{b}_i^*) f(\mathbf{b}_i^*) d\mathbf{b}_i^*} \quad (5)$$

The posterior mean is used to predict the random effects \mathbf{b}_i^* , and the unknown parameters $\boldsymbol{\theta}$ are replaced by their MLEs $\hat{\boldsymbol{\theta}}$. We may use the MCMC to simulate direct draws from the posterior distribution in (5) to obtain the posterior mean.

4 Lack-of-fit test

In this section, we consider the lack-of-fit test. That is, we would like to test if the log-gamma distribution is as good as or better than the normal distribution. Note that the mean and variance of the standard log-gamma distribution W are $E(W) = \psi(\kappa)$ and $\text{Var}(W) = \psi'(\kappa)$, which behave like $\log \kappa$ and κ^{-1} respectively for large κ . Here ψ and ψ' are the digamma and trigamma functions. It has been shown that the transformed log-gamma variate $Z = \kappa^{1/2}(W - \log \kappa)$ converges to the standard normal distribution as $\kappa \rightarrow \infty$ (Bartlett and Kendall, 1946; Prentice, 1974). Therefore, for the lack-of-fit test, we test if $\kappa_0 = \infty$, where κ_0 is the true value of κ .

Let the profile log-likelihood function be $l_p(\kappa)$ for κ . The statistic of the likelihood ratio test $\Lambda(\kappa_0)$ for the hypothesis $H_0: \kappa = \kappa_0$ versus $H_a: \kappa > \kappa_0$ is

$$\Lambda(\kappa_0) = 2l_p(\hat{\kappa}) - 2l_p(\kappa_0),$$

where $\hat{\kappa}$ is the MLE that maximizes $l_p(\kappa)$. For finite κ_0 the distribution of $\Lambda(\kappa_0)$ is asymptotically χ_1^2 under H_0 . A slight technical difficulty arises in testing the normal model, since $\kappa_0 = \infty$ is on the boundary of the parameter space. However, the other parameters in $\boldsymbol{\theta}$ with true values are not on the boundary. Therefore, when $\kappa_0 = \infty$, the asymptotic distribution of $\Lambda(\kappa_0)$ is a 50:50 mixture of χ_0^2 and χ_1^2 ; then for $a \leq 0$, $Pr(\Lambda(\infty) \leq a) = 0.5 + 0.5 Pr(\chi_1^2 \leq a)$ (Self and Liang, 1987).

5 Data analysis

In DIGS, subjects were included if they had a diagnosis of glaucoma or were suspected of having the disease at the baseline. At each visit during the follow-up, the subjects underwent a comprehensive ophthalmologic examination and additional tests designed to evaluate

glaucoma structure and function. They were tested at 6-month intervals. The functional assessment was performed with visual field testing by SAP (the SITA 24-2 test, Humphrey perimeter, Carl-Zeiss Meditec, Dublin, CA). The evaluation of the rates of visual field change during the follow-up was performed using the VFI. The VFI represents the percentage of normal age-corrected visual function, and it is used to calculate rates of progression and the staging of glaucomatous functional damage. The VFI can range from 100% (normal visual field) to 0% (perimetrically blind field). The structural assessment was performed by measuring the thickness of the RNFL using optical coherence tomography (OCT). The RNFL thickness measurements were obtained in a 3.2-mm diameter circle around the optic nerve head. The global average RNFL thickness (calculated as the average of all the RNFL measurements obtained in the circle around the optic nerve) was used in this study.

A total of 1939 repeated measures were collected from each of the 203 patients included in DIGS. The mean (\pm standard deviation) age of the participants was 64.1 ± 12.1 years; 54.5% of the subjects were female and 32.9% were of African ancestry. There was a wide range of disease severities at the baseline, with the SAP mean deviation (MD) values ranging from -19.6 to 2.3 dB and the RNFL thickness values ranging from 40 to $123 \mu\text{m}$.

Because we investigate two types of outcomes simultaneously in this study, we have $p = 2$ in the general model (2). The VFI scores are proportional data, so we apply a logit transformation. The measures of RNFL are regular continuous values. We assume that the disease progression is a linear function of time and that the slope depends on the individual patient. For the i -th patient, the baseline measures of the k -th ($k = 1, 2$) outcome recorded can be expressed as $\beta_{k0} + b_{ik0}$, where β_{k0} is an unknown parameter and b_{ik0} is the random intercept. Several measures are collected at time t_{ijk} , for $j=1, \dots, n_{ik}$. The coefficient β_{k1} is an unknown parameter, b_{ik1} is the random slope for the i -th patient, and the subject-specific progression rate can be expressed as $\beta_{k1} + b_{ik1}t_{ijk}$. Hence, $\mathbf{b}_{i1} = (b_{i10}, b_{i11})'$, $\mathbf{b}_{i2} = (b_{i20}, b_{i21})'$, and $\mathbf{b}_i = (b_{i10}, b_{i11}, b_{i20}, b_{i21})'$. Therefore, we have $q = 4$ in model (3). Furthermore, a random error is associated with each time measurement. Using the vector and matrix notation, the linear-mixed effects model can be written

$$\begin{pmatrix} \mathbf{Y}_{i1} \\ \mathbf{Y}_{i2} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_{i1} & 0 \\ 0 & \mathbf{X}_{i2} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{Z}_{i1} & 0 \\ 0 & \mathbf{Z}_{i2} \end{pmatrix} \begin{pmatrix} \mathbf{b}_{i1} \\ \mathbf{b}_{i2} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_{i1} \\ \boldsymbol{\varepsilon}_{i2} \end{pmatrix}$$

where

$$\mathbf{X}_{ik} = \mathbf{Z}_{ik} = \begin{pmatrix} 1 & t_{i1k} \\ 1 & t_{i2k} \\ \vdots & \vdots \\ 1 & t_{in_{ik}k} \end{pmatrix}, k=1, 2.$$

A preliminary analysis indicated that the intercepts of the VFI scores, the first outcome, show a negatively skewed pattern when only the linear-mixed model for the \mathbf{Y}_{i1} 's is fitted.

Histograms of the regression coefficients for the within-subject regressions of the two outcomes on time are shown in Supporting Information Fig. S1. They not only verify that the distribution of the subject-specific intercepts for VFI is negatively skewed but also clearly indicate the relationship between the two random intercepts and that of the two slopes.

The random intercept of the first outcome b_{i10} in \mathbf{b}_i is the only skewed element. It is in the first position, so there is no need to reorder \mathbf{b}_i to obtain the desired form. That is, $q_0 = 1$ in (3). Then, we define the linear transformation $\mathbf{b}_i = \mathbf{R}\mathbf{s}_i$ with

$$\mathbf{R} = \begin{pmatrix} 1 & r_1 & r_2 & r_3 \\ 0 & 1 & r_4 & r_5 \\ 0 & 0 & 1 & r_6 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\mathbf{s}_i = (s_{i10}, s_{i11}, s_{i20}, s_{i21})'$, with independent components. We model s_{i10} through the log-gamma distribution, and the remaining components in \mathbf{s}_i are multivariate normal. Since the log-gamma distribution is negatively skewed, b_{i10} is negatively skewed also.

Conditioning on the random effects, we assume the responses for a particular individual to be independent and conditionally normal:

$$\mathbf{Y}_i | \mathbf{s}_i \sim N(\mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{R} \mathbf{s}_i, \sum_i), i = 1, \dots, N.$$

Table 1 gives the estimates of the parameters in the log-gamma and normal random effects models. The estimates of the parameters in the log-gamma model are close to those of the normal model. In univariate longitudinal data analysis, deviations from the normality assumption have little impact on the estimation of the fixed effects and the variance components. Based on the parameter estimates in Table 1, the same appears to be true for the multivariate case. We use the multivariate delta method and the asymptotic properties of the MLEs $\hat{\boldsymbol{\theta}}$ to calculate the standard errors of the variance components of \mathbf{b}_i . The introduction of the linear transformation matrix \mathbf{R} does not lead to any difficulties in the computation of the standard errors of the variance component estimates. The computational complexity is similar to that for the setting of the random effects in the normal case; they are usually specified with a nondiagonal covariance matrix. The estimated average regression coefficients for VFI and RNFL are -0.0327 and -0.5310 , respectively. For both outcomes, larger values of the regression coefficients of time indicate a slower deterioration. Therefore, negative slopes indicate disease progression over time.

The profile likelihood functions $l_p(\kappa)$ are derived by maximizing $\log L(\boldsymbol{\theta})$ with κ fixed. The likelihood-ratio statistic $\Lambda(\infty)$ is 2402.4, and the p -value is extremely small: < 0.0001 .

There is strong evidence to reject the null hypothesis that the normal random effects model is equivalent to the log-gamma model, suggesting that the log-gamma model has a better fit to the data.

Table 2 lists the means and standard errors of the $N = 203$ predicted random effects under the log-gamma and normal models, the means of the posterior distributions of \mathbf{b}_i given \mathbf{Y}_i . In the log-gamma model, given estimates of the fixed effects, variance components, and \mathbf{R} , the random effects \mathbf{b}_i were predicted from 18,000 sample points of three chains generated from the posterior distributions $f(\mathbf{b}_i|\mathbf{y}_i)$, with the first 4000 sample points discarded for burn-in. In the normal model, the random effects \mathbf{b}_i were predicted from Empirical Bayes estimates. The distributions of the estimates of the random effects \mathbf{b}_i from the $N = 203$ posterior means under the log-gamma model are shown in Supporting Information Fig. S2. We found that the distribution of the predicted random effects (b_{i10} 's) is negatively skewed, which confirms the observed skewness from the preliminary study. The estimate of the skewness parameter $\kappa = 2.5762$ shows that the estimated distribution of the random slopes is not close to a normal distribution.

The correlations between the two random intercepts, ρ_0 , and between the two random slopes, ρ_1 , of the two outcomes are of most interest. We obtained the correlations of the pairs of random effects and their standard errors from the log-gamma and normal-mixed effects models; they are presented in Table 3. Under the log-gamma model, for the null hypothesis $H_0: \rho_0 = 0$, the Wald test statistic has $z = 21.7790$ with p -value < 0.0001 , which shows that there is a significant positive relationship between the two random intercepts of the regressions of the two test results on time. The positive relationship between the intercepts of the two outcomes is clearly seen in the scatterplot of the intercepts from the single-patient regression analysis, which is shown in the upper-left panel of Fig. 1. The Wald test for the hypothesis regarding the correlation of slopes, $H_0: \rho_1 = 0$, gives $z = 3.1588$ with p -value 0.0016. The intercepts represent the level of disease at the start of the study. The positive correlation coefficient of the two random intercepts indicates that the structural and functional outcomes are consistent with each other. A high baseline in one test will be associated with a high level in the other. Most importantly, the rates of progressive RNFL loss are significantly associated with the rates of functional change in glaucoma, as indicated by the small p -value of the correlation between the two random slopes. Under the normal model, the test for $H_0: \rho_0 = 0$ gives $z = 8.8771$ with p -value $< .0001$, and the test for slopes, $H_0: \rho_1 = 0$, has $z = 1.4845$ with p -value 0.1378. Clearly the model with misspecified distributions of the random effects gives less reliable conclusions on the correlations of the random effects. The normal model underestimates the correlation of the random slopes of the two outcomes, which results in a small value in the test statistic and a large p -value. The log-gamma mixed-effects model gives a better fit to the data, so the results on the correlations of the random effects are more reliable. Hence, the progression rates of function in glaucoma, measured with the VFIs, and the progression rates of structures, measured with the RNFLs, are shown to be positively correlated in glaucoma patients.

6 Discussion

We have proposed a class of log-gamma linear-mixed models for modeling longitudinal data with multiple outcomes and applied it to a longitudinal glaucoma study. To address the main scientific questions of whether the structural and functional measurements are associated and how the progressions of the two outcomes are related, we performed hypothesis tests based on the proposed model. We concluded that the progressions of the structural and

functional outcomes are positively correlated, and the rates of change in the two outcomes are highly correlated.

Besides reflecting the covariance structure of the multivariate responses, the main advantage of the class of log-gamma mixed-effects models lies in its simplicity in accounting for the skewness of the random effects, which results in a more efficient estimation of the parameters. We considered the density functions in the location family of the standard log-gamma distributions with mean zero for the modeling of the skewed random effects. This subfamily of the location-scale family of the log-gamma distributions is sufficient to model the skewness, and it reduces the computational complexity of the search for the MLEs of the marginal log-likelihood function, especially when the dimension of the vector of parameters is high. The expression of the reordered random effects by the product of the linear transformation matrix and the random vector of independent components avoids high-dimensional integration in the marginal log-likelihood function and makes the implementation feasible. The introduction of the linear transformation matrix not only accounts for the correlated random effects, but also ensures that the dimension of the numerical integration of the marginal likelihood is one, since the convolution of two normal distributions is still normal. This new class provides a generalized method for estimating the correlation between two or more slopes. In the case of a single response, the model reduces to that of Zhang et al. (2008). The family of log-gamma distributions possesses the nice property that the limiting distribution is normal, and a lack-of-fit test on the adequacy of the log-gamma distributional assumption of the random effects can be derived. It would be of interest to extend the model to the case of mixed multivariate responses in which the continuous outcome response shows a skewed pattern.

In this bivariate longitudinal data analysis, we observed that the random slopes of time effects from one of the outcomes had a skewed distribution. Modeling this feature using the log-gamma distribution not only led to a better fit than that for the regular multivariate normal-mixed effects models but also led to a different conclusion on the correlation between the two random effects from the two outcomes. This is because the log-gamma model gives more accurate estimates of the random effects by assuming more realistic distributions for the random effects than normal models do. Better accuracy in the estimation of subject-specific effects using the log-gamma model rather than normal models has been observed in univariate longitudinal data analysis (Zhang et al. 2008). Future work will include a theoretical investigation into how misspecifications of the random effects distributions affect the estimation and inference of the correlations of the random effects of multivariate longitudinal data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the editor, an associate editor, and the anonymous referee for constructive comments and valuable suggestions that led to an improvement in the paper. The first and third authors' research was supported by

a start-up grant from Zhejiang University of China and the Natural Sciences and Engineering Research Council of Canada, respectively.

References

- Bartlett MS, Kendall DG. The Statistical analysis of variance-heterogeneity and the logarithmic transformation. *Supplement to the Journal of the Royal Statistical Society*. 1946; 8:128–138.
- Ghosh P, Branco MD, Chakraborty H. Bivariate random effect model using skew-normal distribution with application to HIV-RNA. *Statistics in Medicine*. 2007; 26:1255–1267. [PubMed: 16998836]
- Harwerth RS, Wheat JL, Fredette MJ, et al. Linking structure and function in glaucoma. *Progress in Retinal and Eye Research*. 2010; 29:249–271. [PubMed: 20226873]
- Jiang J. Mixed-effects models with random cluster sizes. *Statistics and Probability Letters*. 2001; 53:201–206.
- Keltner JL, Johnson CA, Anderson DR, et al. The association between glaucomatous visual fields and optic nerve head features in the Ocular Hypertension Treatment Study. *Ophthalmology*. 2006; 113:1603–1612. [PubMed: 16949445]
- Malik R, Swanson WH, Garway-Heath DF. ‘Structure-function relationship’ in glaucoma: past thinking and current concepts. *Clinical and Experimental Ophthalmology*. 2012; 40:369–380. [PubMed: 22339936]
- McCulloch C, Neuhaus J. Prediction of random effects in linear and generalized linear models under model misspecification. *Biometrics*. 2010; 67:270–279. [PubMed: 20528860]
- Medeiros FA, Zangwil LM, Bowd C, et al. The structure and function relationship in glaucoma: implications for detection of progression and measurement of rates of change. *Investigative Ophthalmology and Visual Science*. 2012; 53:6939–6946. [PubMed: 22893677]
- Pocock SJ, Geller NL, Tsiatis AA. The analysis of multiple endpoints in clinical trials. *Biometrics*. 1987; 43:487–498. [PubMed: 3663814]
- Prentice RL. A log gamma model and its maximum likelihood estimation. *Biometrika*. 1974; 61:539–544.
- Reinsel G. Estimation and prediction in a multivariate random-effects generalized linear model. *Journal of the American Statistical Association*. 1984; 79:406–414.
- Roy J, Lin X. Latent variable models for longitudinal data with multiple continuous outcomes. *Biometrics*. 2000; 56:1047–1054. [PubMed: 11129460]
- Ruppert D. Discussion of “Maximization by Parts in Likelihood Inference,” by P. X.-K. Song, Y. Fan, and J. D. Kalbfleisch. *Journal of the American Statistical Association*. 2005; 100:1161–1163.
- Self SG, Liang KY. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*. 1987; 82:605–610.
- Shah A, Laird N, Schoenfeld D. A random-effects model for multiple characteristics with possibly missing data. *Journal of the American Statistical Association*. 1997; 92:775–779.
- Strouthidis NG, Scott A, Peter NM, et al. Optic disc and visual field progression in ocular hypertensive subjects: detection rates, specificity, and agreement. *Investigative Ophthalmology and Visual Science*. 2006; 47:2904–2910. [PubMed: 16799032]
- Thiébaud R, Jacqmin-Gadda H, Chêne G, Leport C, Commenges D. Bivariate linear mixed models using SAS Proc MIXED. *Computer Methods and Programs in Biomedicine*. 2002; 69:249–256. [PubMed: 12204452]
- Weiss L. Asymptotic properties of maximum likelihood estimators in some nonstandard cases, II. *Journal of the American Statistical Association*. 1973; 68:428–430.
- Zhang P, Song P, Qu A, Greene T. Efficient estimation for patient-specific rates of disease progression using nonnormal linear mixed models. *Biometrics*. 2008; 64:29–38. [PubMed: 17501938]

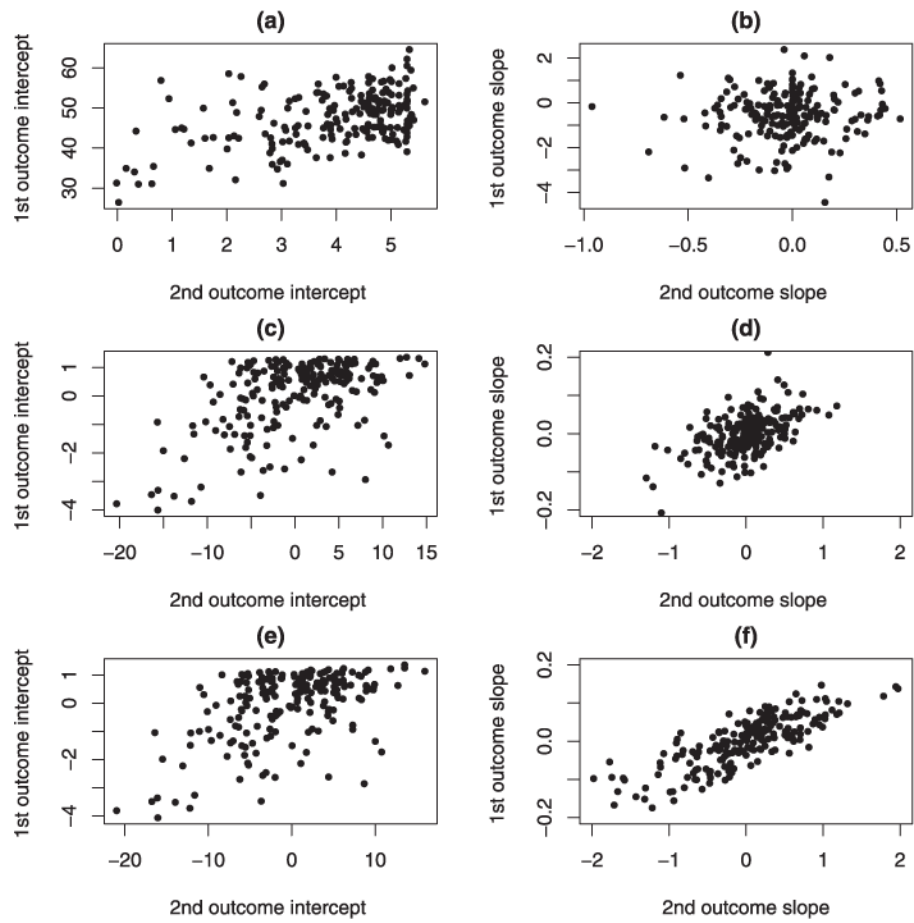


Figure 1. Panels (A) and (B) are scatterplots of the intercepts and slopes for the within-subject regressions of two outcomes on time. Panels (C) and (D) are scatterplots of the empirical estimates of the intercepts and slopes from the posterior distributions of the random effects of the normal random effects model. Panels (E) and (F) are scatterplots of the estimates from the log-gamma random effects model.

Table 1 Estimates of the fixed effects and the variance components of the log-gamma and normal random effects models.

Para.	Log-gamma			Normal		
	Estimate	SE	p-value	Estimate	SE	p-value
β_{10}	4.0425	0.0623	<0.0001	3.9671	0.0919	<0.0001
β_{11}	-0.0327	0.0153	0.0326	-0.0290	0.0148	0.050
β_{20}	47.7384	0.2876	<0.0001	47.4261	0.4630	<0.0001
β_{21}	-0.5310	0.0297	<0.0001	-0.5746	0.0713	<0.0001
$\sigma_{b_{110}}^2$	0.9117	0.0360	<0.0001	1.5998	0.1705	<0.0001
$\sigma_{b_{111}}^2$	0.0031	0.0006	<0.0001	0.0104	0.0040	<0.0001
$\sigma_{b_{120}}^2$	35.4466	3.2269	<0.0001	41.6805	4.3238	<0.0001
$\sigma_{b_{121}}^2$	0.3243	0.0807	<0.0001	0.3948	0.1034	<0.0001
σ_1^2	0.2116	0.0083	<0.0001	0.2396	0.0119	<0.0001
σ_2^2	2.1999	0.1129	<0.0001	2.4707	0.1778	<0.0001

Table 2

Results for the posterior distribution outcomes of random effects under the log-gamma and normal random effects models.

Random effect	Log-gamma		Normal	
	Mean	SE	Mean	SE
b_{i10}	-0.0676	1.2142	-0.0000	1.2444
b_{i11}	0.0006	0.0672	0.0000	0.0538
b_{i20}	-0.3418	6.5187	-0.0002	6.3864
b_{i21}	-0.0186	0.7049	0.0001	0.3978

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

