Harvard University
Harvard University Biostatistics Working Paper Series

Year 2012 *Paper* 145

On a Logistic Mixed Model Formulation of a Quadratic Exponential Model for Correlated Binary Outcomes

Eric J. Tchetgen Tchetgen*

*Harvard University, etchetge@hsph.harvard.edu

http://biostats.bepress.com/harvardbiostat/paper145

Original Article

On a logistic mixed model formulation of a quadratic exponential model for correlated binary outcomes

Eric J. Tchetgen Tchetgen

Departments of Epidemiology and Biostatistics,

Harvard University

Correspondence: Eric J. Tchetgen Tchetgen, Department of Epidemiology, Harvard School of Public Health 677 Huntington Avenue, Boston, MA 02115.

Abstract

In this paper, the author provides a mixed model specification of a general class of quadratic exponential models for correlated binary outcomes, thus effectively establishing that these two seemingly unrelated models, are in fact intimately connected. The connection is particularly fruitful in that it produces an alternative interpretation for the parameters indexing the exponential model and partitions the latter as (i) a vector of subject-specific regression parameters relating covariates to each outcome conditional on a vector of random effects for the cluster; and (ii) a covariance matrix relating the random effects within a cluster. The established equivalence between these two models presents certain computational advantages for modeling and estimating fixed effects and variance components, within the context of complex multilevel data. This is because the exponential model formulation of the logistic mixed effects model readily accommodates, without the need for high dimensional integration, multiple levels of clustering as well as the serial correlations typically present in longitudinal studies. A data example is presented to illustrate the methodology.

KEY WORDS:Logistic mixed model; quadratic exponential family; odds ratio; clustered data; longitudinal data.

1 Introduction

Logistic mixed effects models (Breslow & Clayton 1993) have become a popular approach to modeling correlated binary outcomes. The basic logistic mixed effects model accounts for correlation among clustered observations by incorporating a normally distributed random intercept into the logistic regression of interest. By sharing a single random intercept across observations within a cluster, the basic logistic mixed effects model assumes that the correlation structure of clustered observations is compound symmetric; and the approach can further incorporate random slopes to account for heterogeneity in covariate effects across clusters. However, because of its nonlinear link function, maximum likelihood estimation of the logistic mixed effects model typically requires evaluating for each cluster an integral with respect to the random effects, which in general is not available in closed-form, and is usually evaluated numerically via Gaussian quadrature. An alternative less commonly used approach to modeling correlated binary outcomes entails fitting a quadratic exponential model (Cox, 1972). An advantage with this approach is that the likelihood is available in closed form, and therefore maximum likelihood estimation is relatively straightforward. Although, a key limitation of the approach is the difficulty of directly interpreting the odds ratio parameters relating the covariates to each of the outcomes, mainly because it involves conditioning on the other outcomes within the cluster. As a solution to this problem, Zhao and Prentice (1990) proposed to reparametrize the exponential model in terms of a marginal logistic regression relating each outcome to the covariates, and a correlation matrix for outcomes within a cluster. Fitzmaurice and Laird (1993) similarly proposed a reparametrization in terms of marginal

logistic regression model, but preserved the odds ratio parametrization to encode the within-cluster association of the outcomes; and established an interesting connection between quadratic and more general exponential models and generalized estimating equations for evaluating covariate associations with the marginal risk of the outcome. In the current paper, the author proposes to use the quadratic exponential model in its original parametrization, and provides a logistic mixed effects model interpretation of its canonical parameters, thus effectively establishing that these two seemingly unrelated models, are in fact intimately connected. The connection is particularly fruitful in that it produces an alternative more meaningful interpretation of the parameters indexing the exponential model and the latter are partitioned as (i) a vector of subject-specific regression parameters relating covariates to each outcome conditional on a vector of random effects for the cluster; and (ii) a covariance matrix relating the random effects within a cluster. The established equivalence between these two models presents certain computational advantages for modeling and estimating fixed effects and variance components, within the context of complex multilevel data. This is because the exponential model formulation of the logistic mixed effects model readily accommodates, without the need for high dimensional integration, multiple levels of clustering as well as the serial correlations typically present in longitudinal studies. An appealing feature of the quadratic exponential model is that it reduces to the standard logistic regression in the absence of dependence between the outcomes. The connection to the random effect model thus provides a formulation for the logistic mixed effect model in which the independence logistic model is genuinely embedded, unlike the standard formulation of the logistic mixed model, in which the independence model lies on the boundary of the parameter space. A data example is presented to illustrate the results.

2 Logistic Mixed Model

Let Y_{ij} denote the j^{th} binary outcome, $j = 1, ..., n_i$ from cluster $i, i = 1, ..., N$, and let X_i be a known $n_i \times p$ matrix of covariates constructed so that the jth row X_{ij} is the vector of covariates corresponding to Y_{ij} . The logistic mixed model specifies two components of the model: a model for the vector of outcome variables $Y_i = (Y_{i1}, ..., Y_{in_i})'$ conditional on X_i and unobserved, cluster-specific random effects b_i, and distributional assumptions on b_i. Specifically, we assume Y_{ij} given X_i and b_i are independent random variables with:

$$\text{logit}\mu_{ij} = \text{logit}\Pr\left(Y_{ij} = 1 | X_{ij}, b_i; \beta\right) = \beta_0 + X_{ij}\beta_1 + Z_{ij}b_i \tag{1}$$

where $\beta = \left(\beta_0, \beta_1'\right)$ is a $(p+1)\times 1$ vector of unknown fixed parameters, and Z_{ij} is a known covariate matrix with columns typically a subset of those in X_{ij}. In the special case where $Z_{ij} = 1$, b_i entails a random intercept, and the parameter β_1 captures on the odds ratio scale, the cluster specific effects of X_{ij} on Y_{ij} given b_i, so that $\beta = 0$ encodes the independence of Y_{ij} and X_{ij} given b_i. The mixture component of the likelihood is defined by $b_i \sim f_b(b_i|\ \Sigma_i(\delta))$, where $f_b(\cdot|\Sigma_i(\delta))$ is a smooth joint density of the vector of random effects such that the usual regularity conditions for the standard maximum likelihood theory hold (Cox & Hinkley 1974). This density is parameterized by the covariance matrix $\Sigma_i(\delta)$ indexed by an unknown parameter δ. A common specification of the random intercept model takes $b_i \sim N(0, \delta^2)$; although more flexible specifications have also been considered in the literature, see Molenberghs and Verbeke (2005). For inference, the marginal likelihood

$$\prod_{i=1}^{N} L_{M,i}(\beta, \delta) = \prod_{i=1}^{N} \int f\left(\widetilde{Y}_i, b_i | \widetilde{X}_i; \beta, \Sigma_i(\delta)\right) db_i$$

is constructed under the additional assumption that the elements of Y_i are conditionally independent given b_i :

$$\begin{aligned}
&\prod_{i=1}^{N} L_{M,i}\left(\beta,\delta\right) \\
&=\prod_{i=1}^{N}\int f\left(Y_i|X_i,b_i;\beta\right) f\left(b_i|X_i;\Sigma_i\left(\delta\right)\right) db_i \\
&=\prod_{i=1}^{N}\int \prod_{j=1}^{n_i} \mu_{ij}\left(\beta\right)^{Y_{ij}} \left(1-\mu_{ij}\left(\beta\right)\right)^{1-Y_{ij}} f_b(b_i|\Sigma_i\left(\delta\right))db_i
\end{aligned} \tag{2}$$

and maximum marginal likelihood estimates of (β,δ) are obtained upon maximizing $\sum_{i=1}^{n}\log L_{M,i}\left(\beta,\delta\right)$ and standard errors are constructed by inverting the information matrix of the marginal likelihood.

3 Quadratic exponential model

The quadratic exponential model assumes the density of $[Y_i|X_i]$ is of the form

$$\exp\left\{Y_i^T\Omega_i Y_i + C_i\right\} \tag{3}$$

where Ω_i is an $n_i \times n_i$ matrix function of X_i, and C_i is the normalizing constant

$$C_i = -\log\sum_{y\in\{0,1\}^{n_i}}\exp\{y^T\Omega_i y\}$$

Because Y_{ij} is binary, the jth diagonal entry of Ω_i, may be interpreted as the log odds of Y_{ij}, i.e.

$$\Omega_i^{(j,j)} = \Omega_i^{(j,j)}\left(X_i\right) = \log\{\Pr(Y_{i,j}=1|Y_{i,-j}=0,X_i)/\Pr(Y_{i,j}=0|Y_{i,-j}=0,X_i)\} \tag{4}$$

where $Y_{i,-j}$ is the subvector of Y_i obtained by deleting $Y_{i,j}$. The off-diagonal entry

$$\Omega_i^{(j,j')} = \Omega_i^{(j,j')}(X_i) = \log \frac{\Pr(Y_{i,j}=1|Y_{i,j'}=1, Y_{i,-(j,j')}, X_i)/\Pr(Y_{i,j}=0|Y_{i,j'}=1, Y_{i,-(j,j')}, X_i)}{\Pr(Y_{i,j}=1|Y_{i,j'}=0, Y_{i,-(j,j')}, X_i)/\Pr(Y_{i,j}=0|Y_{i,j'}=0, Y_{i,-(j,j')}, X_i)} \tag{5}$$

is the log odds ratio relating Y_{ij} and $Y_{ij'}$ conditional on $\left(Y_{i,-(j,j')}, X_i\right)$, $j \neq j'$. Thus, it is straightforward to verify that $\Omega_i^{(j,j')} = 0$ for all $j \neq j'$ encodes the null hypothesis that $Y_{i,j}$ and $Y_{i,j}$ are independent; while $\Omega_i^{(j,j')} \neq 0$ implies that there is dependence of the outcomes measured within a cluster. In principle, one could model the effect of $X_{i,j}$ on the risk of $Y_{i,j}$ by specifying a logistic regression model for $\Omega_i^{(j,j)}(X_i)$, say

$$\begin{aligned}
&\Omega_i^{(j,j)}(X_i) - \Omega_i^{(j,j)}(0) \\
&= \log\left\{\frac{\Pr(Y_{i,j}=1|Y_{i,-j}=0, X_i)/\Pr(Y_{i,j}=0|Y_{i,-j}=0, X_i)}{\Pr(Y_{i,j}=1|Y_{i,-j}=0, X_i=0)/\Pr(Y_{i,j}=0|Y_{i,-j}=0, X_i=0)}\right\} \\
&= X_{i,j}\psi
\end{aligned} \tag{6}$$

which encodes the simplifying assumption that $\Omega_i^{(j,j)}$ only depends on X_i through $X_{i,j}$. Unfortunately, the parameter ψ is generally difficult to interpret because it captures the association between $X_{i,j}$ and $Y_{i,j}$ upon conditioning on the other outcomes within the cluster; i.e. conditional on $Y_{i,-j} = 0$, an association measure seldom of primary interest. As mentioned in the introduction, in an effort to resolve this difficulty, Zhao and Prentice (1990) and Fitzmaurice and Laird (1993) respectively proposed a reparametrization of the density (3) that entails instead of the logistic model (6), specifying a marginal logistic regression for $E(Y_{i,j}|X_{i,j})$ that relates $X_{i,j}$ to $Y_{i,j}$, $j = 1, ..n_i$. In the next section, instead of reparametrizing the quadratic exponential model as proposed by Zhao and Prentice (1990) and Fitzmaurice and Laird (1993), we propose a reparametrization of the logistic mixed model (2), such that the regression parameters $\psi = \beta_1$ and therefore ψ can be

interpreted as a cluster-specific effect of $X_{i,j}$ on $Y_{i,j}$.

4 A mixed model interpretation of the quadratic exponential model

4.1 General formulation

To state the main result, we consider a reparametrization of the joint density $f(Y_i, b_i|X_i)$ of the outcome and the random effects conditional on X_i, where for the moment we make no parametric assumption about the functional form of this density. We proceed as in Tchetgen Tchetgen et al (2010) and note that the joint density of $[Y_i, b_i|X_i]$ can generally be written:

$$f(Y_i, b_i|X_i) = \frac{g_0\left(Y_i|X_i\right) OR(Y_i, b_i|X_i) h_0(b_i|X_i)}{\sum_{y\in\{0,1\}^{n_i}} \int g_0\left(y|X_i\right) OR(y, \widetilde{b}|X_i) h_0(\widetilde{b}|X_i) d\widetilde{b}}$$

where $g_0\left(Y_i|X_i\right) = f(Y_i|X_i, b_i = 0), h_0(b_i|X_i) = f(b_i|X_i, Y_i = 0)$ and

$$OR(Y_i, b_i|X_i) = \frac{f(Y_i, b_i|X_i) f(Y_i = 0, b_i = 0|X_i)}{f(Y_i = 0, b_i|X_i) f(Y_i, b_i = 0|X_i)}$$

is the conditional odds ratio function relating b_i and Y_i within levels of X_i; and assuming

$$\sum_{y\in\{0,1\}^{n_i}} \int g_0\left(y|X_i\right) OR(y, \widetilde{b}|X_i) h_0(\widetilde{b}|X_i) d\widetilde{b} < \infty.$$

The above expression effectively replaces the marginal density of the random effects $f(b_i|X_i)$ with the conditional density $f(b_i|X_i, Y_i = 0)$ in parametrizing the joint distribution of $[Y_i, b_i|X_i]$. However, as we show next, this reparametrization is perfectly compatible with a logistic mixed model,

but the assumptions about the random effects distribution are inherently different under the reparametrization. In fact, to be consistent with the logistic mixed model (1), let

$$g_0(Y_i|X_i, b_i = 0; \beta) = \prod_{j=1}^{n_i} \mu_{ij}^0(\beta)^{Y_{ij}} \left(1 - \mu_{ij}^0(\beta)\right)^{1-Y_{ij}}$$

where $\mu_{ij}^0(\beta) = \text{logitPr}(Y_{ij} = 1|X_{ij}, b_i = 0; \beta)$ is given by (1); and

$$\log OR(Y_i, b_i|X_i) = \sum_{j=1}^{n_i} Z_{ij} b_i Y_{ij}$$

Then, it is easy to verify that this specification recovers:

$$f(Y_i|X_i, b_i) = \prod_{j=1}^{n_i} \mu_{ij}(\beta)^{Y_{ij}} (1 - \mu_{ij}(\beta))^{1-Y_{ij}}$$

To proceed with inference under this reparametrization, we assume that the random effect density $h_0(b_i|X_i)$ is multivariate normal:

$$[b_i|Y_i = 0, X_i] \sim MVN(0, \Sigma_i(\gamma))$$

with covariance matrix indexed by an unknown parameter γ. Let $f(Y_i, b_i|X_i; \beta, \gamma)$ denote the joint density under this specification; then a straightforward application of the moment generating function of the multivariate normal distribution produces the following equivalence between the marginal likelihood of $[Y_i|X_i]$ for the reparametrized logistic mixed model, and the quadratic exponential model:

$$\int f(Y_i, b_i|X_i; \beta, \gamma) db_i = \exp\left\{Y_i^T \widetilde{\Omega}_i(\beta, \gamma) Y_i + \widetilde{C}_i(\beta, \gamma)\right\}$$

where

$$\widetilde{\Omega}_i^{(j,j)}(\beta,\gamma) = \beta_0 + X_{ij}\beta_1 + Z_{i,j}\Sigma_i(\gamma) Z_{i,j}^T/2,$$

$$\widetilde{\Omega}_i^{(j,j')}(\beta,\gamma) = \; Z_{i,j}\Sigma_i(\gamma) Z_{i,j'}^T/2,$$

and $\widetilde{C}_i(\beta,\gamma) = \; -\log\sum_{y\in\{0,1\}^{n_i}}\exp\{y^T\widetilde{\Omega}_i(\beta,\gamma)\, y\}$. Estimation of (β,γ) then entails maximizing the log likelihood $\sum_i Y_i^T\widetilde{\Omega}_i(\beta,\gamma) Y_i + \widetilde{C}_i(\beta,\gamma)$, and variance estimates of the maximum likelihood estimator can be obtained by inverting the corresponding information matrix.

4.2 Random intercept logistic models

An important special case is the random intercept model where $Z_{i,j} = 1$ and $\Sigma_i(\gamma) = \gamma^2$, then the formulae in the previous display yield

$$\begin{aligned}
&\int f(Y_i, b_i|X_i;\psi,\gamma)db_i \\
&= \exp\left\{Y_i^T\widetilde{\Omega}_i Y_i + \widetilde{C}_i\right\} \\
&= \frac{\exp\left\{\sum_{j=1}^{n_i}(\beta_0 + X_{ij}\beta_1 + \gamma^2/2) Y_{ij} + \sum_{1\le j\neq j'\le n_i}\gamma^2 Y_{ij}Y_{ij'}\right\}}{\sum_{y\in\{0,1\}^{n_i}}\exp\left\{\sum_{j=1}^{n_i}(\beta_0 + X_{ij}\beta_1 + \gamma^2/2) y_j + \sum_{1\le j\neq j'\le n_i}\gamma^2 y_j y_{j'}\right\}}
\end{aligned} \tag{7}$$

As we show next, the connection between the logistic mixed model and the quadratic exponential model also facilitates estimation of more general mixed models.

For instance, consider the more general random intercept logistic model:

$$\text{logit}\mu_{ij} = \text{logit}\Pr(Y_{ij} = 1|X_{ij}, b_i;\beta) = \beta_0 + X_{ij}\beta_1 + b_{i,j} \tag{8}$$

which allows each observation within a cluster to have a separate intercept, β_1 is an observation specific covariate effect, and $b_i = (b_{i,1}, ...b_{i,n_i})$. Let $\gamma_{j,j'}$ denote the covariance $\text{Cov}(b_{i,j}, b_{i,j'}|Y_i =$

$0, Z_i)$, and suppose that $[b_{i,j} \,|Y_i = 0, Z_i] \sim N(0, \gamma_{jj})$. Then, a similar derivation as above gives:

$$\int f(Y_i, b_i|X_i; \beta, \gamma) db_i$$
$$= \frac{\exp\left\{\sum_{j=1}^{n_i} (\beta_0^* + X_{ij}\beta_1\ ) Y_{ij} + \sum_{1\le j\neq j'\le n_i} \gamma_{j,j'} Y_{ij} Y_{ij'}\ \right\}}{\sum_{y\in\{0,1\}^{n_i}} \exp\left\{\sum_{j=1}^{n_i} (\beta_0^* + X_{ij}\beta_1)\, y_j + \sum_{1\le j\neq j'\le n_i}^{2} \gamma_{j,j'} y_j y_{j'}\ \right\}}$$

where here, we note that $\beta_0^* = \beta_0 + \gamma_{j,j}/2$ and β_1 identified regression parameters, but β_0 and $\gamma_{j,j}$ are not separately identifiable. In other words, the variance of $b_{i,j}$ is not identified but the covariance components $\gamma_{j,j'}$ are identified for $j \neq j'$. The above model recovers the shared random effect model, i.e the standard random intercept model, upon specifying $\gamma_{j,j'} = \gamma_{j,j} = \gamma^2$, in which case of course, $\gamma_{j,j'}$ becomes identified. A potential limitation of the standard random intercept model is that it produces a compound symmetric correlation structure in which the outcomes are restricted to be positively correlated. The connection to the quadratic model motivates alternative simple formulations of the random intercept model in which the outcome are not a priori restricted to be positively correlated, as illustrated in the data example in Section 5. One such formulation might specify $\gamma_{j,j} = \gamma_0 \geq 0$ and $\gamma_{j,j'} = Cov(b_{i,j}, b_{i,j'}|Y_i = 0) = \gamma_1$ constant in j,j', such that γ_1 is unrestricted and thus can take on a positive or a negative value; therefore accommodating an exchangeable covariance structure for possibly negatively correlated binary outcomes. In principle, by virtue of γ_1 being unrestricted, the above model could be used to construct a standard likelihood ratio test of the null hypothesis that the outcomes are independent, *i.e.*that the covariance components $\gamma_{j,j'} = 0$ for all $j \neq j'$.

In a longitudinal setting where j indexes time, $\gamma_{j,j'}$ could easily be modelled to reflect the typical serial correlation structures encountered in such settings; for instance by assuming $\gamma_{j,j'} = \gamma_0 \exp^{-\gamma_1|t_i - t_j|^2}$so that the correlation between observations is weaker the further apart they are,

with γ_0, γ_1 unknown parameters. A further generalization may be made to handle a multilevel setting in which longitudinal measurements are made on clustered binary outcomes $\{Y_{i,j,t} : i, j, t\}$, for unit j at time t, within cluster i. Such a multilevel setting is easily captured for instance, by assuming that the random intercepts within a cluster have a covariance structure given by

$$\gamma_{j,j',t_1,t_2} = Cov(b_{i,j,t_1}, b_{i,j',t_2}|Y_i = 0) = \gamma_2 + \gamma_0 \exp^{-\gamma_1|t_1-t_2|^2}$$

consisting of a serial correlation component reflecting the longitudinal nature of the data and an exchangeable correlation component reflecting clustering at each occasion.

4.3 Random effect logistic models

In this section, we briefly consider another generalization of the simple random intercept models of the previous section and we additionally allow for a single random slope:

$$\text{logit}\mu_{ij} = \text{logit}\Pr\left(Y_{ij} = 1|X_{ij}, b_i; \beta\right) = \beta_0 + X_{ij}\beta_1 + b_{1,i} + Z_{ij}b_{2,i}$$

where Z_{ij} is a scalar variable contained in X_{ij}, $\left[(b_{1i}, b_{2,i})^T |Y_i = 0, X_i\right]$ is multivariate normal with mean zero and $Var(b_{1,i}|Y_i = 0, X_i) = \gamma_1, Var(b_{2,i}|Y_i = 0, X_i) = \gamma_1, Cov(b_{1,i}, b_{2,i}|Y_i = 0, X_i) = \gamma_{1,2}$. Thus, we obtain using the results from the previous sections:

$$\begin{aligned}&\int f(Y_i, b_i|X_i; \beta, \gamma)db_i \\ &= \exp\left\{Y_i^T\Omega_i^{\dagger}\left(\beta, \gamma\right)Y_i + C_i^{\dagger}\left(\beta, \gamma\right)\right\}\end{aligned}$$

where

$$\Omega_i^{\dagger(j,j)}\left(\beta, \gamma\right) = \beta_0 + X_{ij}\beta_1 + \gamma_1/2 + \gamma_2 Z_{i,j}^2 + \gamma_{1,2}Z_{i,j}/2,$$

$$\Omega_i^{\dagger(j,j')}(\beta,\gamma) = \gamma_1/2 + \gamma_{12}\left(Z_{i,j} + Z_{i,j'}\right)/2 + \gamma_2 Z_{i,j'}/2$$

In principle, a more general model along the lines of the observation specific random intercept model could similarly allow $b_{1,i,j}$ to vary across observations within a cluster, details are omitted but are easily deduced from the presentation.

5 A data application

Fitzmaurice, Laird, & Ware (2004) used a logistic generalized linear mixed model with random intercepts to analyze data from a longitudinal clinical trial examining the effects of hormonal contraceptives in women. In the trial, contracepting women received four successive injections of either 100 mg or 150 mg of depot-medroxyprogesterone acetate at 0, 90, 180, and 270 days after randomization, with this dosage remaining constant for each subject over the course of the study. There was also a final follow-up visit one year after the first injection. The analysis, which was based on N = 1151 women, focused on the within subject effects of time on the binary outcome of whether a woman experienced amenorrhea in the four successive three-month intervals, and whether this trend in risk varied according to dosage. Let $Y_{ij} = 1$ if woman i,$i = 1, ..., 1151$, experienced amenorrhea in the jth injection interval, $j = 1, ..., 4$,and $Y_{ij} = 0$ otherwise. Fitzmaurice, Laird & Ware (2004) considered the model

$$\text{logit}\{\Pr(Y_{ij} = 1|b_i)\} = \beta_1 + \beta_2 time_{ij} + \beta_3 time_{ij}^2 + \beta_4 dose_i \times time_{ij} + \beta_5 dose_i \times time_{ij}^2 + b_i,$$

where $time_{ij} = 1, 2, 3, 4$ for the four consecutive 90-day injection intervals and $\text{dose}_i = 1$ if subject i is randomized to 150mg of depot-medroxyprogesterone acetate and $\text{dose}_i = 0$ otherwise. The

model specifies a quadratic within-subject effect of time, with this trend differing according to the dosage received. Because of randomization, the model does not include a main effect of drug, which corresponds to assuming that no differences exist between the two drug groups at baseline. In this model, dose_i is a between-subject effect and time_{ij} is a within-subject effect. Fitzmaurice, Laird & Ware (2004) completed the specification of the model by assuming that

$$b_i \sim N(0, \gamma^2),$$

i.e. the correlation structure within an individual is compound symmetric, and thus the outcomes are positively correlated.

Insert Table 1 here.

Table 1 presents the parameter estimates for $(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ and corresponding standard errors from the fit of this model. Results suggest that there is a significant effect of dose on the trend for the risk of amenorrhea, and that there is a large amount of heterogeneity in the baseline risk among subjects. Table 1 also presents the results from a quadratic exponential model fit, corresponding to the following more general logistic random intercept model :

$$\text{logit}\{\Pr(Y_{ij} = 1|b_i)\} = \beta_1 + \beta_2 time_{ij} + \beta_3 time_{ij}^2 + \beta_4 dose_i \times time_{ij} + \beta_5 dose_i \times time_{ij}^2 + b_{i,j},$$

$$j = 1, ..., 4$$

$$b_i = (b_{i1}, b_{i2,} b_{i3}, b_{i4})^T \, | Y_i = 0, X_i \sim MVN(0, \Sigma(\gamma))$$

$\Sigma(\gamma)$ with entries $\gamma^{j,j} = \widetilde{\gamma}^2$ and unstructured off-diagonal elements $\gamma^{j,j'}$

The above mixed model is more general than the model considered by Fitzmaurice, Laird, & Ware (2004) in that similar to model (8), each person-time observation has a unique intercept, this allows the correlation structure relating observations within an individual to remain unstructured. Upon specifying this more flexible random intercept model, we note that the magnitude of the effects of the intervention is essentially halved, and only the effect on a linear trend β_4 remains statistically significant. Upon inspecting the covariance components $\gamma_{j,j'}$ of $\Sigma(\gamma)$, it is quite striking that the covariance structure of the random intercepts do not appear to follow any specific standard pattern, and the results provides evidence that consecutive outcomes are positively related, but outcomes two or more occasions apart appear to be negatively correlated. These results further suggest that the simple random intercept model fit by Fitzmaurice, Laird, & Ware (2004) may not be entirely appropriate for these data.

6 Conclusion

This paper unveils a simple relation between a logistic mixed model and a general class of quadratic exponential models for correlated binary outcomes, two modeling approaches that have until now, thought to be unrelated. As formally established, and illustrated in a data application, the equivalence between these two models is computationally advantageous for modeling and estimating fixed effects and variance components, particularly in the presence of complex correlation structures; this is primarily because the formulation of the logistic mixed model as a quadratic exponential model permits inferences based on a simple closed-form likelihood. We anticipate that the reparametrization used in this paper to reveal the equivalence between these two models might also be useful in other logistic latent variable models, such as for for instance, a logistic regression with additive measurement error-in-variables.

References

[1] Breslow, N. E. & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. J. Am. Statist. Assoc. 88, 9–25.

[2] Cox, D. R. (1972). The analysis of multivariate binary data. Appl Statist. 21, 113-20.

[3] Fitzmaurice, G. M. and Laird, N. M. (1993). A likelihood-based method for analysing longitudinal binary responses Biometrika. 80. 141-151.

[4] Cox, D. R. and Hinkley, D. V. (1974). Theoretical Statistics. London: Chapman & Hall.

[5] Molenberghs G, Verbeke G. (2005) Models for Discrete Longitudinal Data. Springer-Verlag New York.

[6] Tchetgen Tchetgen E, Robins JM, Rotnitzky A. (2010). On doubly robust estimation in a semiparametric odds ratio model. Biometrika 97(1):171-180.

[7] Zhao, L. P. and Prentice, R. L. (1990). Correlated binary regression using a quadratic exponential model. Biometrika 77, 642-8.

TABLE 1.Logistic-normal mixed model and quadratic exponential MLEs for the Amenorrhea Data.

Variable	Logistic-normal (SE)	Quadratic (SE)
Intercept	-3.8057 (0.3050)	-3.6786 (0.4694)
time_{ij}	1.1332 (0.2682)	0.4470 (0.3800)
time_{ij}^2	-0.0419 (0.0548)	0.0308 (0.0700)
$\text{dose}_i \times \text{time}_{ij}$	0.5644 (0.1922)	0.2234 (0.1076)
$\text{dose}_i \times \text{time}_{ij}^2$	-0.1095 (0.0496)	-0.0588 (0.0327)
γ^2	5.0646 (0.5840)	-
$\gamma_{1,2}$	-	3.8792 (0.2603)
$\gamma_{1,3}$	-	-1.8180 (0.4911)
$\gamma_{1,4}$	-	-1.0928 (0.4834)
$\gamma_{2,3}$	-	0.3644 (0.3588)
$\gamma_{2,4}$	-	-3.2339 (0.4753)
$\gamma_{3,4}$	-	7.5456 (0.3787)