Application of the Disposition Model to Breast Cancer Data

Reginald N. O. Odai[1], Wolfgang Urfer[1], Joachim Kötting[1]
and George E. Bonney[2]

[1]Department of Statistics, University of Dortmund, D-44221 Dortmund, Germany
[2]National Human Genome Center, Howard University, Washington, D. C., U.S.A.

Summary

In this paper, we have presented the second level nesting of Bonney's disposition model (Bonney, 1998) and examined the implications of higher level nesting of the disposition model in relation to the dimension of the parameter space. We have also compared the performance of the disposition model with Cox's regression model (Cox, 1972). It has been observed that the disposition model has a very large number of unknown parameters, and is therefore limited by the method of estimation used. In the case of the maximum likelihood method, reasonable estimates are obtained if the number of parameters in the model is at most nine. This corresponds to about four to seven covariates. Since each covariate in Cox's model provides a parameter, it is possible to include more covariates in the regression analysis. On the other hand, as opposed to Cox's model, the disposition model is fitted with parameters to capture aggregation in families, if there should be any. The choice of a particular model should therefore depend on the available data set and the purpose of the statistical analysis.

Key words: Second level nesting; Proportional hazards model; Quadratic exponential form; Partial likelihood; Familial aggregation; Second-order methods; Marginal models; Conditional models.

1 Introduction

The outcomes of family members are correlated because they share common risks. Thus standard methods of epidemiology, which assume independence of outcomes, are unsuitable for the analysis of family data. Many models have been proposed to incorporate dependence within families. Cox (1972) reviewed several methods that had been proposed for the analysis of multivariate binary data and outlined some new proposals. He suggested the use of logistic representations, in which the joint response probability is a quadratic exponential form, as the simplest, most flexible, and in many ways the most important models. In the paper 'Partial likelihood', Cox (1975) gave a definition of partial likelihood which generalises the ideas of conditional and marginal likelihood. Here, he transformed the random variable Y into a sequence $\{X_j, S_j\}$, $j = 1,...,m$, and decomposed the full likelihood of the sequence into two products, the second product being the partial likelihood based on S in the sequence $\{X_j, S_j\}$. He pointed out that the partial likelihood is especially useful when it is appreciably simpler than the full likelihood. This is the situation when constructive procedures for finding useful partial likelihoods are provided, so that the partial likelihood involves only the parameters of interest and not nuisance parameters. To support this point, he made mention of the failure of the method of maximum likelihood as a general technique, especially in the sampling theory and pure likelihood approaches, due to excessive nuisance parameters, and hence the need to reduce dimensions. Care should however be taken to ensure that all or nearly all the relevant information is contained in the partial likelihood.

Connolly and Liang (1988) introduced the conditional logistic regression models for correlated binary data which are most useful when the dependence among observations is of main interest (such as in family data). Although the estimating functions are easily computed and have high efficiency compared to the computationally intensive maximum likelihood approach, more work is needed to determine the form of the weights used for the estimating functions $U(\beta,\theta)$. Zhao and Prentice (1990) reparameterised probability distribution of the model advocated by Cox (1972) in terms of marginal parameters of ready interpretation. Since this approach yields models with very complicated marginal response probabilities and pairwise correlations, they suggested the transformations of the canonical parameters (θ_k, λ_k), $k = 1,...,K$, to response means ($\mu_k = \mu_k(\beta)$) and covariances ($\sigma_k = \sigma_k(\beta,\alpha)$), where β and α are parameter vectors. Scoring estimating functions can then be used to evaluate mean and correlation parameters under the quadratic exponential family. Liang, Zeger and Qaqish (1992) presented a model for correlated binary data, in which the marginal expectation of each binary variable as well as the association between pairs of outcomes are modelled separately in terms of explanatory variables. With examples, they described some drawbacks of conditional models, especially in situations where observations are missing or cluster sizes differ. On the other hand, the marginal model is reproducible, since the marginal distribution of any proper subset $(Y_1,\ldots,Y_n)$ is of the same form. Hence the situation where a subset of the cluster $(Y_1,\ldots,Y_n)$ is missing causes no problem. Carey, Zeger and Diggle (1993) proposed the use of odds ratios to measure association among responses. The approach, which alternates between two steps, estimates the association parameters by modelling the conditional distributions of one response given another. The alternating logistic regression avoids the computational burdens encountered in many problems, and its estimates are reasonably efficient relative to solutions of second-order methods. Odai et al. (2002) discussed the use of the correlated Weibull regression model for the analysis of multivariate binary data. The results have shown that the model provides feasible means of analysing family data.

In this paper, the implications of higher level nesting in relation to the dimension of the parameter space are examined. Since in real life, levels of nesting higher than two serve no practical purpose, we will limit our work to second level nesting. Section 2 briefly reviews Cox's regression model (Cox, 1972) for the analysis of failure data when explanatory variables are available. Section 3 contains a detailed treatment of the second level nesting of the disposition model. The estimation procedure will be described in Section 4. The method is illustrated with breast cancer data in Section 5, and is followed by discussion.

2 Cox's regression model

The Cox model (also known as the proportional hazards model) is a model that can be used for the analysis of failure data when explanatory variables are available.

2.1 The model

Let $h(t;x)$ be the hazard rate at time t for an individual with risk vector $x^T = (x_1,...,x_p)$. Cox (1972) specified his model as follows:

$$h(t;x) = h_0(t)\exp(\beta^T x), \tag{2.1.1}$$

where $h_0(t)$ is an arbitrary baseline hazard rate and $\beta^T = (\beta_1,...,\beta_p)$ is a vector of unknown parameters.

The above model is often called a proportional hazards model because, the ratio of the hazard rates of two individuals with covariate values x and x' can be expressed as

$$\frac{h(t;x)}{h(t;x')} = \exp\left[\sum_{k=1}^{p} \beta_k (x_k - x_k^{'})\right], \tag{2.1.2}$$

which is a constant (see, for example, Klein and Moeschberger, 1997). This indicates that the hazard rates are proportional. The quantity (2.1.2), called the relative risk (hazard ratio), gives the factor by which the risk of an individual with covariate x is increased in comparison to an individual with risk factor x'.

2.2 Parameter estimation

In order to estimate the parameters in Cox's model with the maximum likelihood method, the baseline hazard, $h_0(t)$, must be specified. To deal with this situation, Cox exploited the definition of partial likelihood. Specifically, he considered the baseline hazard, $h_0(t)$, as a nuisance parameter function and concentrated mainly on the regression parameters.

Let $t_{(1)} < t_{(2)} < \ldots < t_{(n)}$ denote the ordered event times and define the risk set at time $t_{(i)}$, $R(t_{(i)})$, $i = 1,\ldots,n$, as the set of all individuals who are still under study at a time just prior to $t_{(i)}$. Further, let x_j denote the value of x for the jth individual, and $x_{(i)}$ the value for the individual failing at time $t_{(i)}$, $i = 1,\ldots,n$. Then, Cox (1972) gave the partial likelihood based on the hazard function specified by (2.1.1) as

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(\beta^T x_{(i)})}{\sum_{j \in R(t_{(i)})} \exp(\beta^T x_j)}. \tag{2.2.1}$$

It should be noted that the numerator of the likelihood in (2.2.1) depends only on information from the individual who experiences the event, whereas the denominator utilises information about all individuals who have not yet experienced the event (Klein and Moeschberger, 1997).

Direct calculation from the log-likelihood gives the score equation

$$U(\beta) = \sum_{i=1}^{n} x_{(i)} - \sum_{i=1}^{n} \frac{\sum_{j \in R(t_{(i)})} x_j \exp(\beta^T x_j)}{\sum_{j \in R(t_{(i)})} \exp(\beta^T x_j)}, \tag{2.2.2}$$

from which we obtain the Hessian matrix

$$H(\beta) = A_{(i)}(\beta)A_{(i)}^T(\beta) - \sum_{i=1}^{n} \frac{\sum_{j \in R(t_{(i)})} x_j x_j^T \exp(\beta^T x_j)}{\sum_{j \in R(t_{(i)})} \exp(\beta^T x_j)}, \tag{2.2.3}$$

where $A_{(i)} = \frac{\sum_{j \in R(t_{(i)})} x_j \exp(\beta^T x_j)}{\sum_{j \in R(t_{(i)})} \exp(\beta^T x_j)}$, $i = 1, \ldots, n$.

The Fisher information matrix is given by

$$I(\beta) = \sum_{i=1}^{n} \frac{\sum_{j \in R(t_{(i)})} x_j x_j^T \exp(\beta^T x_j)}{\sum_{j \in R(t_{(i)})} \exp(\beta^T x_j)} - \sum_{i=1}^{n} \left[\frac{\sum_{j \in R(t_{(i)})} x_j \exp(\beta^T x_j)}{\sum_{j \in R(t_{(i)})} \exp(\beta^T x_j)} \right] \left[\frac{\sum_{j \in R(t_{(i)})} x_j^T \exp(\beta^T x_j)}{\sum_{j \in R(t_{(i)})} \exp(\beta^T x_j)} \right], \tag{2.2.4}$$

(Klein and Moeschberger, 1997). Cox (1975) has shown that the usual maximum likelihood properties hold for estimates and tests based on partial likelihoods.

3 Second level nesting

Consider a binary outcome Y = 1 or 0, with q_0 primary-group-specific covariates (i.e., cluster-specific covariates), $Z_0^T = (Z_{01}, \ldots, Z_{0q_0})$, q_i secondary-group-specific covariates (i.e., subgroup-specific covariates), $Z_i^T = (Z_{i1}, \ldots, Z_{iq_i})$, $i = 1, \ldots, m$, q_{ij} tertiary-group-specific covariates, $Z_{ij}^T = (Z_{ij1}, \ldots, Z_{ijq_{ij}})$, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, and p unit-specific covariates,

$X_{ijh}^T = (X_{ijh1},\ldots,X_{ijhp})$, $i=1,\ldots,m$, $j=1,\ldots,n_i$, $h=1,\ldots,n_{ij}$, measured on several units. Four types of dispositions are considered here: the group (cluster) disposition, δ_0, which is determined by the group-specific covariates, Z_0, the subgroup disposition, δ_i, $i=1,\ldots,m$, which is determined by the group-specific covariates, Z_0, and the subgroup-specific covariates, Z_i, $i=1,\ldots,m$, the tertiary-group disposition, δ_{ij}, which is determined by the primary-group-specific covariates, Z_0, the secondary-group-specific covariates, Z_i, and the tertiary-group-specific covariates, Z_{ij}, and the unit disposition, δ_{ijh}, $i=1,\ldots,m$, $j=1,\ldots,n_i$, $h=1,\ldots,n_{ij}$, which is determined by the primary-group-specific covariates, Z_0, the secondary-group-specific covariates, Z_i, the tertiary-group-specific covariates, Z_{ij}, and the unit-specific covariates, X_{ijh}, $i=1,\ldots,m$, $j=1,\ldots,n_i$, $h=1,\ldots,n_{ij}$.

We define δ_0, δ_i and δ_{ij} as follows:

$$\delta_0 = \frac{\mu_0}{\alpha_0}, \tag{3.1}$$

$$\delta_i = \frac{\mu_i}{\alpha_i}, \tag{3.2}$$

$i=1,\ldots,m$, and

$$\delta_{ij} = \frac{\mu_{ij}}{\alpha_{ij}}, \tag{3.3}$$

$i=1,\ldots,m$, $j=1,\ldots,n_i$, where μ_0 is the primary group baseline disposition under no aggregation, μ_i is the secondary group baseline disposition under no aggregation,

μ_{ij} is the tertiary group baseline disposition under no aggregation, α_0 is the relative disposition with respect to the primary group, α_i is the relative disposition with respect to the secondary group and α_{ij} is the relative disposition with respect to the tertiary group.

The logit of the unit disposition is decomposed as

$$\log\frac{\delta_{ijh}}{1-\delta_{ijh}} = M_0(Z_0) + D_0(Z_0) + M_i(Z_i) + D_i(Z_i) + M_{ij}(Z_{ij}) + D_{ij}(Z_{ij}) + W_{ijh}(X_{ijh})$$

$$=: \theta_{ijh}, \tag{3.4}$$

$i = 1,\ldots,m$, $j = 1,\ldots,n_i$, $h = 1,\ldots,n_{ij}$, where

$$M_0(Z_0) = \log\frac{\mu_0}{1-\mu_0} \tag{3.5}$$

is the cluster logit mean risk,

$$D_0(Z_0) = \log\frac{\delta_0}{1-\delta_0} - \log\frac{\mu_0}{1-\mu_0} \tag{3.6}$$

is the excess cluster logit disposition due to dependence among members of the group,

$$M_i(Z_i) = \log\frac{\mu_i}{1-\mu_i} - \log\frac{\delta_0}{1-\delta_0}, \tag{3.7}$$

$i = 1,\ldots,m$, is the excess on the logit scale of the mean risk in secondary group i above that due to the cluster disposition,

$$D_i(Z_i) = \log\frac{\delta_i}{1-\delta_i} - \log\frac{\mu_i}{1-\mu_i}, \tag{3.8}$$

$i = 1,\dots,m$, is the excess on the logit scale of the secondary group i disposition that cannot be explained by the overall primary group disposition and the differences in μ_i,

$$M_{ij}(Z_{ij}) = \log\frac{\mu_{ij}}{1-\mu_{ij}} - \log\frac{\delta_i}{1-\delta_i}, \tag{3.9}$$

$i = 1,\dots,m$, $j = 1,\dots,n_i$, is the excess on the logit scale of the mean risk in the tertiary group j above that due to the secondary group disposition,

$$D_{ij}(Z_{ij}) = \log\frac{\delta_{ij}}{1-\delta_{ij}} - \log\frac{\mu_{ij}}{1-\mu_{ij}}, \tag{3.10}$$

$i = 1,\dots,m$, $j = 1,\dots,n_i$, is the excess on the logit scale of the tertiary group disposition that cannot be explained by the overall cluster disposition, the subgroup disposition and the differences in μ_{ij}, and

$$W_{ijh}(X_{ijh}), \tag{3.11}$$

$i = 1,\dots,m$, $j = 1,\dots,n_i$, $h = 1,\dots,n_{ij}$, is a function of the unit-specific covariates.

From (3.5)-(3.10), we have

$$\mu_0 = \frac{1}{1+\exp\{-[M_0(Z_0)]\}}, \quad \delta_0 = \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)]\}},$$

$$\mu_i = \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)]\}}, \ i = 1,\dots,m,$$

$$\delta_i = \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)]\}}, \ i = 1,\dots,m,$$

$$\mu_{ij} = \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)+M_{ij}(Z_{ij})]\}},$$

$i = 1,\dots,m$, $j = 1,\dots,n_i$, and

$$\delta_{ij} = \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)+M_{ij}(Z_{ij})+D_{ij}(Z_{ij})]\}},$$

$i = 1,\dots,m$, $j = 1,\dots,n_i$. (3.12)

Hence,

$$\alpha_0 = \frac{\mu_0}{\delta_0} = \frac{1+\exp\{-[M_0(Z_0)+D_0(Z_0)]\}}{1+\exp\{-[M_0(Z_0)]\}}, \quad (3.13)$$

$$\alpha_i = \frac{\mu_i}{\delta_i} = \frac{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)]\}}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)]\}}, \quad (3.14)$$

$i = 1,\dots,m$,

$$\alpha_{ij} = \frac{\mu_{ij}}{\delta_{ij}} = \frac{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)+M_{ij}(Z_{ij})+D_{ij}(Z_{ij})]\}}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)+M_{ij}(Z_{ij})]\}},$$

$$i = 1,\dots,m,\ \ j = 1,\dots,n_i, \qquad (3.15)$$

and

$$\delta_{ijh} = \frac{1}{1+\exp\{-\theta_{ijh}\}}$$

$$= \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)+M_{ij}(Z_{ij})+D_{ij}(Z_{ij})+W_{ijh}(X_{ijh})]\}},$$

$$i = 1,\dots,m,\ \ j = 1,\dots,n_i,\ \ h = 1,\dots,n_{ij}. \qquad (3.16)$$

The joint probability for a cluster is

$$P(Y_{111} = y_{111},\dots,Y_{mn_in_{ijh}} = y_{mn_in_{ijh}}) = (1-\alpha_0)\prod_{i=1}^{m}\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})$$

$$+\ \alpha_0\prod_{i=1}^{m}\left\{(1-\alpha_i)\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})+\alpha_i\prod_{j=1}^{n_i}\left\{(1-\alpha_{ij})\prod_{h=1}^{n_{ij}}(1-y_{ijh})\right.\right.$$

$$\left.\left.+\ \alpha_{ij}\prod_{h=1}^{n_{ij}}\delta_{ijh}^{y_{ijh}}(1-\delta_{ijh})^{1-y_{ijh}}\right\}\right\}. \qquad (3.17)$$

The following parameterisations are considered:

$$M_0(Z_0) = \xi_{00} + \xi_{01}Z_{01} + \dots + \xi_{0q_0}Z_{0q_0}, \qquad (3.18)$$

$$D_0(Z_0) = \gamma_{00} + \gamma_{01}Z_{01} + \dots + \gamma_{0q_0}Z_{0q_0}, \qquad (3.19)$$

$$M_i(Z_i) = \xi_1 Z_{i1} + \dots + \xi_q Z_{iq_i}, \qquad (3.20)$$

$i = 1,\dots,m$,

$$D_i(Z_i) = \gamma_1 Z_{i1} + \ldots + \gamma_q Z_{iq_i}, \quad (3.21)$$

$i = 1,\ldots,m$,

$$M_{ij}(Z_{ij}) = \xi_{11} Z_{ij1} + \ldots + \xi_{1q} Z_{ijq_{ij}}, \quad (3.22)$$

$i = 1,\ldots,m$, $j = 1,\ldots,n_i$, and

$$D_{ij}(Z_{ij}) = \gamma_{11} Z_{ij1} + \ldots + \gamma_{1q} Z_{ijq_{ij}}, \quad (3.23)$$

$i = 1,\ldots,m$, $j = 1,\ldots,n_i$.

The set of parameters to be determined in the model is therefore

$$\lambda = (\xi,\gamma,\beta) = (\xi_{00},\ldots,\xi_{0q_0},\xi_1,\ldots,\xi_{q_i},\xi_{11},\ldots,\xi_{1q_{ij}},\gamma_{00},\ldots,\gamma_{0q_0},\gamma_1,\ldots,\gamma_{q_i},\gamma_{11},\ldots,\gamma_{1q_{ij}},\beta_1,\ldots,\beta_p).$$

4 Parameter estimation for the second level nesting

Denote the likelihood function in (3.17) by $L_k(\lambda\,|\,y)$, $k = 1,\ldots,K$:

$$\begin{aligned} L_k(\lambda|y) &= (1-\alpha_0)\prod_{i=1}^{m}\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh}) + \alpha_0 L_{\pi i} \\ &= \prod_{i=1}^{m}\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh}) + \alpha_0\left[L_{\pi i} - \prod_{i=1}^{m}\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})\right], \end{aligned} \quad (4.1)$$

where $L_{\pi i} = \prod_{i=1}^{m} L_i$, $L_i = (1-\alpha_i)\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh}) + \alpha_i L_{\pi j}$, $L_{\pi j} = \prod_{j=1}^{n_i} L_j$,

$L_j = (1-\alpha_{ij})\prod_{h=1}^{n_{ij}}(1-y_{ijh}) + \alpha_{ij} L_{\pi h}$, $L_{\pi h} = \prod_{h=1}^{n_{ij}} L_h$, $L_h = \delta_{ijh}^{y_{ijh}}(1-\delta_{ijh})^{1-y_{ijh}}$ and

$$\delta_{ijh} = \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)+M_{ij}(Z_{ij})+D_{ij}(Z_{ij})+W_{ijh}(X_{ijh})]\}},$$

$$W_{ijh}(X_{ijh}) = (1-\exp(\beta_1 X_{ijh1} + \ldots + \beta_p X_{ijhp})),\ i=1,\ldots,m,\ j=1,\ldots,n_i,\ h=1,\ldots,n_{ij}.$$

The corresponding score function is

$$U_k(\lambda \mid y) = A_k(\lambda)\alpha_0^* + B_k(\lambda)\left[\sum_{i=1}^{m} U_i\right], \tag{4.2}$$

$k=1,\ldots,K$, where $\alpha_0^* = -(1-\delta_0)\frac{\delta}{\delta\lambda}D_0(Z_0) + \delta_0(1-\alpha_0)\frac{\delta}{\delta\lambda}M_0(Z_0)$,

$$A_k(\lambda) = \frac{\alpha_0\left[L_{\pi i} - \prod_i^m \prod_j^{n_i} \prod_h^{n_{ij}} (1-y_{ijh})\right]}{L_k},\ k=1,\ldots,K,\ B_k(\lambda) = \frac{\alpha_0 L_{\pi i}}{L_k},\ k=1,\ldots,K,$$

$$U_i(\lambda \mid y) = A_i(\lambda)\alpha_i^* + B_i(\lambda)\left[\sum_{j=1}^{n_i} U_j\right],\ i=1,\ldots,m,$$

$$\alpha_i^* = -(1-\delta_i)\frac{\delta}{\delta\lambda}D_i(Z_i) + \delta_i(1-\alpha_i)\frac{\delta}{\delta\lambda}[M_0(Z_0) + D_0(Z_0) + M_i(Z_i)],\ i=1,\ldots,m,$$

$$A_i(\lambda) = \frac{\alpha_i\left[L_{\pi j} - \prod_{j=1}^{n_i} \prod_{h=1}^{n_{ij}} (1-y_{ijh})\right]}{L_i},\ i=1,\ldots,m,\ B_i(\lambda) = \frac{\alpha_i L_{\pi j}}{L_i},\ i=1,\ldots,m,$$

$$U_j(\lambda \mid y) = A_j(\lambda)\alpha_{ij}^* + B_j(\lambda)\left[\sum_{h=1}^{n_{ij}} U_h\right],\ j=1,\ldots,n_i,$$

$$\alpha_{ij}^* = -\ (1-\delta_{ij})\frac{\delta}{\delta\lambda}D_{ij}(Z_{ij})$$

$$+\ \delta_{ij}(1-\alpha_{ij})\frac{\delta}{\delta\lambda}\{[M_0(Z_0) + D_0(Z_0) + M_i(Z_i) + D_i(Z_i) + M_{ij}(Z_{ij})]\},$$

$$A_j(\lambda) = \frac{\alpha_{ij}\left[L_{\pi h} - \prod_{h=1}^{n_{ij}} (1-y_{ijh})\right]}{L_j},\ j=1,\ldots,n_i,\ B_j(\lambda) = \frac{\alpha_{ij} L_{\pi h}}{L_j},\ j=1,\ldots,n_i,$$

$$U_h(\lambda\,|\,y) = (y_{ijh} - \delta_{ijh})\theta_{ijh}^{(1)}$$

$$= (y_{ijh} - \delta_{ijh})\frac{\delta}{\delta\lambda}[M_0(Z_0) + D_0(Z_0) + M_i(Z_i) + D_i(Z_i) + M_{ij}(Z_{ij}) + D_{ij}(Z_{ij}) + W_{ijh}(X_{ijh})]$$

$$= (y_{ijh} - \delta_{ijh})\begin{pmatrix} \frac{\delta}{\delta\xi_0}M_0(Z_0) \\ \frac{\delta}{\delta\gamma_0}D_0(Z_0) \\ \frac{\delta}{\delta\xi_i}M_i(Z_i) \\ \frac{\delta}{\delta\gamma_i}D_i(Z_i) \\ \frac{\delta}{\delta\xi_{ij}}M_{ij}(Z_{ij}) \\ \frac{\delta}{\delta\gamma_{ij}}D_{ij}(Z_{ij}) \\ \frac{\delta}{\delta\beta}W_{ij}(X_{ij}) \end{pmatrix} = (y_{ijh} - \delta_{ijh})\begin{pmatrix} Z_0^{'} \\ Z_0^{'} \\ Z_i^{'} \\ Z_i^{'} \\ Z_{ij}^{'} \\ Z_{ij}^{'} \\ -X_{ijh}\exp(\beta^T X_{ijh}) \end{pmatrix},$$

$Z_0^T = (1, Z_{01}, Z_{02}, \ldots, Z_{0q_0})$, $Z_i^T = (Z_{i1}, \ldots, Z_{iq_i})$, $Z_{ij}^T = (Z_{ij1}, \ldots, Z_{ijq_{ij}})$, $\beta^T = (\beta_1, \ldots, \beta_p)$ and $X_{ijh}^T = (X_{ijh1}, \ldots, X_{ijhp})$, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$, $h = 1, \ldots, n_{ij}$.

The Hessian matrix is given by

$$H_k(\lambda) = B_k\left[\sum_{i=1}^{m} H_i(\lambda)\right] + \frac{\prod_{i=1}^{m}\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_k}A_k\alpha_0^*\alpha_0^{*T} + \frac{\prod_{i=1}^{m}\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_k}B_k(1-\alpha_0)\left[\sum_{i=1}^{m}U_i\right]\left[\sum_{i=1}^{m}U_i\right]^T$$

$$+ \frac{\prod_{i=1}^{m}\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_k}B_k\left\{\left[\sum_{i=1}^{m}U_i\right]\alpha_0^{*T} + \alpha_0^*\left[\sum_{i=1}^{m}U_i\right]^T\right\} + A_k\frac{\delta}{\delta\lambda^T}\alpha_0^*, \qquad (4.3)$$

k = 1,…,K, where

$$H_i = \frac{\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_i} A_i \alpha_i^* \alpha_i^{*T} + \frac{\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_i} B_i \left[\alpha_i^* \left[\sum_{j=1}^{n_i} U_j\right]^T + \left[\sum_{j=1}^{n_i} U_j\right] \alpha_i^{*T} \right]$$

$$+ \frac{(1-\alpha_i)\prod_{j=1}^{n_i}\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_i} B_i \left[\sum_{j=1}^{n_i} U_j\right]\left[\sum_{j=1}^{n_i} U_j\right]^T + B_i \left[\sum_{j=1}^{n_i} H_j\right] + A_i \frac{\delta}{\delta\lambda^T}\alpha_i^*,$$

$$H_j = \frac{\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_j} A_j \alpha_{ij}^* \alpha_{ij}^{*T} + \frac{\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_j} B_j \left[\alpha_{ij}^* \left[\sum_{h=1}^{n_{ij}} U_h\right]^T + \left[\sum_{h=1}^{n_{ij}} U_h\right] \alpha_{ij}^{*T} \right]$$

$$+ \frac{(1-\alpha_{ij})\prod_{h=1}^{n_{ij}}(1-y_{ijh})}{L_j} B_j \left[\sum_{h=1}^{n_{ij}} U_h\right]\left[\sum_{h=1}^{n_{ij}} U_h\right]^T + B_j \left[\sum_{h=1}^{n_{ij}} H_h\right] + A_j \frac{\delta}{\delta\lambda^T}\alpha_{ij}^*$$

and

$$H_h(\lambda) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -(y_{ijh}-\delta_{ijh})X_{ijh}X_{ijh}^T \exp(2\beta^T X_{ijh}) \end{pmatrix}$$

$$-\delta_{ijh}(1-\delta_{ijh}) \begin{pmatrix} Z_0^{'}Z_0^{'T} & Z_0^{'}Z_0^{'T} & Z_0^{'}Z_i^T & Z_0^{'}Z_i^T & Z_0^{'}Z_{ij}^T & Z_0^{'}Z_{ij}^T & -Z_0^{'}w^T \\ Z_0^{'}Z_0^{'T} & Z_0^{'}Z_0^{'T} & Z_0^{'}Z_i^T & Z_0^{'}Z_i^T & Z_0^{'}Z_{ij}^T & Z_0^{'}Z_{ij}^T & -Z_0^{'}w^T \\ Z_iZ_0^{'T} & Z_iZ_0^{'T} & Z_iZ_i^T & Z_iZ_i^T & Z_iZ_{ij}^T & Z_iZ_{ij}^T & -Z_iw^T \\ Z_iZ_0^{'T} & Z_iZ_0^{'T} & Z_iZ_i^T & Z_iZ_i^T & Z_iZ_{ij}^T & Z_iZ_{ij}^T & -Z_iw^T \\ Z_{ij}Z_0^{'T} & Z_{ij}Z_0^{'T} & Z_{ij}Z_i^T & Z_{ij}Z_i^T & Z_{ij}Z_{ij}^T & Z_{ij}Z_{ij}^T & -Z_{ij}w^T \\ Z_{ij}Z_0^{'T} & Z_{ij}Z_0^{'T} & Z_{ij}Z_i^T & Z_{ij}Z_i^T & Z_{ij}Z_{ij}^T & Z_{ij}Z_{ij}^T & -Z_{ij}w^T \\ -wZ_0^{'T} & -wZ_0^{'T} & -wZ_i^T & -wZ_i^T & -wZ_{ij}^T & -wZ_{ij}^T & X_{ijh}X_{ijh}^T\exp(2\beta^TX_{ijh}) \end{pmatrix},$$

$w = [X_{ijh}\exp(\beta^T X_{ijh})]$, $i = 1,\dots,m$, $j = 1,\dots,n_i$, $h = 1,\dots,n_{ij}$.

If we should use the correlated logistic regression model, we would have the following corresponding expressions for δ_{ijh}, $U_h(\lambda \mid y)$ and $H_h(\lambda)$:

$$\delta_{ijh} = \frac{1}{1+\exp\{-[M_0(Z_0)+D_0(Z_0)+M_i(Z_i)+D_i(Z_i)+M_{ij}(Z_{ij})+D_{ij}(Z_{ij})+W_{ijh}(X_{ijh})]\}} ,$$

$W_{ijh}(X_{ijh}) = \beta_1 X_{ijh1} + \ldots + \beta_p X_{ijhp}$, $i = 1,\ldots,m$, $j = 1,\ldots,n_i$, $h = 1,\ldots,n_{ij}$,

$$U_h(\lambda \mid y) = = (y_{ijh} - \delta_{ijh}) \begin{pmatrix} Z_0' \\ Z_0' \\ Z_i' \\ Z_i' \\ Z_{ij}' \\ Z_{ij}' \\ X_{ijh} \end{pmatrix},$$

$Z_0^T = (1, Z_{01}, Z_{02}, \ldots, Z_{0q_0})$, $Z_i^T = (Z_{i1}, \ldots, Z_{iq_i})$, $Z_{ij}^T = (Z_{ij1}, \ldots, Z_{ijq_{ij}})$ and $X_{ijh}^T = (X_{ijh1}, \ldots, X_{ijhp})$, $i = 1,\ldots,m$, $j = 1,\ldots,n_i$, $h = 1,\ldots,n_{ij}$, and

$$H_h(\lambda) = -\delta_{ijh}(1-\delta_{ijh}) \begin{pmatrix} Z_0'Z_0^T & Z_0'Z_0^T & Z_0'Z_i^T & Z_0'Z_i^T & Z_0'Z_{ij}^T & Z_0'Z_{ij}^T & -Z_0'X_{ijh}^T \\ Z_0'Z_0^T & Z_0'Z_0^T & Z_0'Z_i^T & Z_0'Z_i^T & Z_0'Z_{ij}^T & Z_0'Z_{ij}^T & -Z_0'X_{ijh}^T \\ Z_iZ_0^T & Z_iZ_0^T & Z_iZ_i^T & Z_iZ_i^T & Z_iZ_{ij}^T & Z_iZ_{ij}^T & -Z_iX_{ijh}^T \\ Z_iZ_0^T & Z_iZ_0^T & Z_iZ_i^T & Z_iZ_i^T & Z_iZ_{ij}^T & Z_iZ_{ij}^T & -Z_iX_{ijh}^T \\ Z_{ij}Z_0^T & Z_{ij}Z_0^T & Z_{ij}Z_i^T & Z_{ij}Z_i^T & Z_{ij}Z_{ij}^T & Z_{ij}Z_{ij}^T & -Z_{ij}X_{ijh}^T \\ Z_{ij}Z_0^T & Z_{ij}Z_0^T & Z_{ij}Z_i^T & Z_{ij}Z_i^T & Z_{ij}Z_{ij}^T & Z_{ij}Z_{ij}^T & -Z_{ij}X_{ijh}^T \\ X_{ijh}Z_0^T & X_{ijh}Z_0^T & X_{ijh}Z_i^T & X_{ijh}Z_i^T & X_{ijh}Z_{ij}^T & X_{ijh}Z_{ij}^T & X_{ijh}X_{ijh}^T \end{pmatrix},$$

$i = 1,\ldots,m$, $j = 1,\ldots,n_i$, $h = 1,\ldots,n_{ij}$.

The Fisher information matrix for the second level nesting is

$$
\begin{aligned}
I_k(\lambda) \ = \ & \alpha_0 \sum_{i=1}^{m} I_i(\lambda) - A_k^* \alpha_0^* \alpha_0^{*T} - B_k^* (1-\alpha_0) \left[\sum_{i=1}^{m} U_i^* \right] \left[\sum_{i=1}^{m} U_i^* \right]^T \\
& - \ B_k^* \left\{ \alpha_0^* \left[\sum_{i=1}^{m} U_i^* \right]^T + \left[\sum_{i=1}^{m} U_i^* \right] \alpha_0^{*T} \right\},
\end{aligned} \tag{4.4}
$$

$k = 1,\ldots,K$, where

$$
\begin{aligned}
I_i(\lambda) = & \alpha_i \sum_{j=1}^{n_i} I_j(\lambda) - A_i^* \alpha_i^* \alpha_i^{*T} - B_i^* (1-\alpha_i) \left[\sum_{j=1}^{m} U_j^* \right] \left[\sum_{j=1}^{m} U_j^* \right]^T \\
& - \ B_i^* \left[\alpha_i^* \left[\sum_{j=1}^{n_i} U_j^* \right]^T + \left[\sum_{j=1}^{n_i} U_j^* \right] \alpha_i^{*T} \right],
\end{aligned}
$$

$$
\begin{aligned}
I_j(\lambda) = & \alpha_{ij} \sum_{h=1}^{n_{ij}} I_h(\lambda) - A_j^* \alpha_{ij}^* \alpha_{ij}^{*T} - B_j^* (1-\alpha_{ij}) \left[\sum_{h=1}^{n_{ij}} U_h^* \right] \left[\sum_{h=1}^{n_{ij}} U_h^* \right]^T \\
& - \ B_j^* \left[\alpha_{ij}^* \left[\sum_{h=1}^{n_{ij}} U_h^* \right]^T + \left[\sum_{h=1}^{n_{ij}} U_h^* \right] \alpha_{ij}^{*T} \right],
\end{aligned}
$$

$$
\begin{aligned}
I_h(\lambda) &= \delta_{ijh}(1-\delta_{ijh}) \theta_{ijh}^{(1)} \theta_{ijh}^{(1)T} \\
&= \delta_{ijh}(1-\delta_{ijh}) \begin{pmatrix}
Z_0' Z_0^{'T} & Z_0' Z_0^{'T} & Z_0' Z_i^T & Z_0' Z_i^T & Z_0' Z_{ij}^T & Z_0' Z_{ij}^T & -Z_0' w^T \\
Z_0' Z_0^{'T} & Z_0' Z_0^{'T} & Z_0' Z_i^T & Z_0' Z_i^T & Z_0' Z_{ij}^T & Z_0' Z_{ij}^T & -Z_0' w^T \\
Z_i Z_0^{'T} & Z_i Z_0^{'T} & Z_i Z_i^T & Z_i Z_i^T & Z_i Z_{ij}^T & Z_i Z_{ij}^T & -Z_i w^T \\
Z_i Z_0^{'T} & Z_i Z_0^{'T} & Z_i Z_i^T & Z_i Z_i^T & Z_i Z_{ij}^T & Z_i Z_{ij}^T & -Z_i w^T \\
Z_{ij} Z_0^{'T} & Z_{ij} Z_0^{'T} & Z_{ij} Z_i^T & Z_{ij} Z_i^T & Z_{ij} Z_{ij}^T & Z_{ij} Z_{ij}^T & -Z_{ij} w^T \\
Z_{ij} Z_0^{'T} & Z_{ij} Z_0^{'T} & Z_{ij} Z_i^T & Z_{ij} Z_i^T & Z_{ij} Z_{ij}^T & Z_{ij} Z_{ij}^T & -Z_{ij} w^T \\
-wZ_0^{'T} & -wZ_0^{'T} & -wZ_i^T & -wZ_i^T & -wZ_{ij}^T & -wZ_{ij}^T & X_{ijh} X_{ijh}^T \exp(2\beta^T X_{ijh})
\end{pmatrix},
\end{aligned}
$$

$w = [X_{ijh} \exp(\beta^T X_{ijh})]$, $i = 1,\ldots,m$, $j = 1,\ldots,n_i$, $h = 1,\ldots,n_{ij}$, and A_k^*, A_i^*, A_j^*, B_k^*, B_i^*, B_j^*, U_i^*, U_j^* and U_h^* are the resulting values of A_k, A_i, A_j, B_k, B_i, B_j, U_i, U_j and U_h evaluated at $y = 0$.

5 Illustration

Data are available on 240 families with breast cancer in the national database and at the Howard University, Washington, D. C., U.S.A.. The data set comprises family data and epidemiology data. The variables to be assessed are annual household income (hinc), age at time of examination (ageat), obesity, and tumour of the breast other than breast cancer (tumour). Family-specific data consist of hinc in thousands (<5, 5-15, 15-25, 25-35, 35-50, 50+), whereas subject-specific data consist of ageat in years and obesity (0 - not obese; 1 - obese), and unit or breast-specific data consist of tumour (0 – absence; 1 - presence). The response variable indicates whether or not a breast is affected with breast cancer. This is coded as 0 for unaffected and 1 for affected. Two levels of nesting exist in these data: two breasts are nested within each subject and subjects are nested within families (compare with the second example of Qaqish and Liang, 1992). The objective of the analysis is to assess the presence of familial aggregation of breast cancer.

Model for the second level nesting:

The variables in the model are hinc (Z_{01}), ageat (Z_{ij1}), obesity (Z_{ij2}) and tumour (X_{ijh1}). That is, we have one group-specific covariate, two subject-specific covariates and one unit-specific covariate. There are no subgroup-specific covariates. The linear models in Equations (3.18) – (3.23) can therefore be specified as follows:
$M_0(Z_0) = \xi_{00} + \xi_{01}Z_{01}$, $D_0(Z_0) = \gamma_{00} + \gamma_{01}Z_{01}$, $M_i(Z_i) = 0$, $D_i(Z_i) = 0$,
$M_{ij}(Z_{ij}) = \xi_{11}Z_{ij1} + \xi_{12}Z_{ij2}$, and $D_{ij}(Z_{ij}) = \gamma_{11}Z_{ij1} + \gamma_{12}Z_{ij2}$.

For the function that describes the effects of the unit-specific covariate, we have
$W_{ijh}(X_{ijh}) = 1 - \exp(\beta X_{ijh1})$ for the correlated Weibull regression model and
$W_{ijh}(X_{ijh}) = \beta X_{ijh1}$ for the correlated logistic regression model. The set of parameters to be determined in the model is

$$\lambda = (\xi, \gamma, \beta) = (\xi_{00}, \xi_{01}, \xi_{11}, \xi_{12}, \gamma_{00}, \gamma_{01}, \gamma_{11}, \gamma_{12}, \beta).$$

Parameter estimates and standard deviations of the estimates, along with Wald statistics are given in Table 5.1 for the correlated Weibull and the correlated logistic regression models. The function of the individual-specific covariates, $W_{ijh}(X_{ijh})$, is equal to zero, since no breast has a primary tumour other than breast cancer. Hence, the estimates for both regression models are the same. The parameter β is fixed for computational reasons. The covariates of positive (negative) coefficients increase (decrease) the probability for breast cancer.

For the 1-parameter Wald's tests, the null hypothesis that $\gamma_j = 0$ is rejected for γ_{00}, γ_{01} and γ_{11}. This is an indication of the existence of familial aggregation of breast cancer. On the other hand, the null hypothesis of $\gamma_j = 0$ for γ_{12} cannot be rejected at the level $\alpha = 0.05$. Hence, obesity does not affect the familial aggregation of breast cancer.

Table 5.1: Parameter estimates, standard deviations of the estimates and Wald statistics for the correlated Weibull and the correlated logistic regression models

Variable	Parameter	Correlated Weibull and logistic regression models		
		Parameter estimate	Standard deviation	Wald statistic
constant	ξ_{00}	-3.9134	0.2759	14.1841
hinc	ξ_{01}	0.2791	0.0189	14.7672
ageat	ξ_{11}	0.0250	0.0034	7.3529
obesity	ξ_{12}	0.1275	0.1118	1.1404
constant	γ_{00}	-0.6350	0.1530	4.1503
hinc	γ_{01}	-0.0477	0.0105	4.5429
ageat	γ_{11}	-0.0290	0.0131	2.2137
obesity	γ_{12}	0.7268	0.9904	0.7338
tumour	β	---	---	---

Critical value for the rejection of the null hypothesis: $u_{0.975} = 1.96$.

For the global tests, the hypotheses to be tested are $H_0: \gamma = 0$ and $H_1: \gamma \neq 0$, where $\gamma^T = (\gamma_{00}, \gamma_{01}, \gamma_{11}, \gamma_{12})$.

Let $\log L_0$ = the maximised log-likelihood from which γ is omitted,

$\log L_1$ = the full log-likelihood and

$\chi_4^2 = 9.4877$, the critical value for rejection of the null hypothesis.

Then, the likelihood ratio statistic for the correlated Weibull and the correlated logistic regression models is LR = -2[$\log L_0 - \log L_1$] = -2[-536.1829 – (-466.6963)] = 138.9732. Thus, significant familial aggregation is observed for both regression models (see, for example, Wilks, 1938).

The maximum likelihood estimate of γ is $\hat{\gamma} = \begin{pmatrix} 6.2049\text{x}10^{-2} \\ -2.1115\text{x}10^{-2} \\ -4.5258\text{x}10^{-3} \\ 8.1942\text{x}10^{-1} \end{pmatrix}$ and the estimated variance-covariance matrix is

$$\hat{\text{var}}(\hat{\gamma}) = \begin{pmatrix} 4.31\text{x}10^{-1} & -1.11\text{x}10^{-4} & -4.72\text{x}10^{-7} & 6.86\text{x}10^{-4} \\ -1.11\text{x}10^{-4} & 3.13\text{x}10^{-5} & 6.86\text{x}10^{-4} & -2.50\text{x}10^{-4} \\ -4.72\text{x}10^{-7} & 6.86\text{x}10^{-4} & 3.95\text{x}10^{-6} & -1.77\text{x}10^{-4} \\ 6.86\text{x}10^{-4} & -2.50\text{x}10^{-4} & -1.77\text{x}10^{-4} & 3.08\text{x}10^{-2} \end{pmatrix}.$$

The Wald statistic for $H_0 : \gamma = 0$ has a value of

$W = (\hat{\gamma} - 0)^T I_{\gamma\gamma}(\hat{\gamma}, \hat{\gamma})(\hat{\gamma} - 0) = \hat{\gamma}^T [\hat{\text{var}}(\hat{\gamma})]^{-1} \hat{\gamma} = 32.2734$ (see Garthwaite et al., 1995; Bickel and Doksum, 1977). Since the Wald statistic is large, the null hypothesis will be rejected. The conclusion is that there is aggregation of breast cancer in families.

Table 5.2 presents estimates of the parameters obtained by fitting Cox’s model, with standard deviations and Wald statistics for testing effects.

Table 5.2: Parameter estimates, standard deviations of the estimates and Wald statistics resulting from Cox's model

Variable	Parameter	Parameter estimate	Standard deviation	Wald statistic
hinc	β_1	-0.0159	0.0196	0.8112
ageat	β_2	-0.0050	0.0020	2.5000
obesity	β_3	0.0441	0.0652	0.6764
tumour	β_4	---	---	---

Critical value for the rejection of the null hypothesis: $u_{0.975} = 1.96$.

The hypothesis to be tested is $H_0 : \beta_j = 0$ versus $H_1 : \beta_j \neq 0$. From Table 5.2, the covariate ageat is the only significant factor. The covariates hinc and obesity produce non-significant effects, since the values of their Wald statistics are less than the critical value, $u_{0.975} = 1.96$. For the global null hypothesis $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$, we obtain 6.9773 for the likelihood ratio statistic and 6.9874 for the Wald statistic, both values indicating non-significance when compared to a chi-square distribution with three degrees of freedom (i.e., $\chi^2_{3,0.95} = 7.8147$).

6 Discussion

In this paper, we have discussed the second level nesting of the disposition model and its estimating procedure for the analysis of correlated binary data. We have also investigated the problems associated with estimation as the level of nesting gets deeper, and compared the performance of the nested disposition model with Cox's model (Cox, 1972). The main disadvantage of the disposition model is that, with the exception of the unit-specific covariates, each covariate in the model produces two parameters. This results in the following problems:

(1) The effect of a covariate can have different interpretations. For instance, in Table 5.1 the covariate hinc increases the cluster logit mean risk, $M_0(Z_0) = \xi_{00} + \xi_{01} Z_{01}$, whereas the same covariate decreases the excess cluster logit disposition due to dependence among members of

the group, $D_0(Z_0) = \gamma_{00} + \gamma_{01} Z_{01}$. Thus, the same variable hinc gives two opposing effects with regard to the probability for breast cancer,

$$\delta_{ijh} = \frac{1}{1 + \exp\{-[M_0(Z_0) + D_0(Z_0) + M_{ij}(Z_{ij}) + D_{ij}(Z_{ij})]\}}.$$

(2) The number of covariates that can be included in the model is seriously limited. We recall that in a previous work, we could estimate up to nine parameters from five covariates, using the maximum likelihood method (see Table 6.2.1, Odai et al., 2002). An attempt to estimate more than nine parameters from five covariates (the fifth covariate finally excluded from the analysis) in the present work resulted in over-identified parameters (i.e., parameters estimated in two or more linearly independent ways).

The disposition model however has the advantage that aggregations in families, due to common shared risks, and response probabilities can jointly be modelled.

Acknowledgements

This work was funded by the Studentensekretariat, University of Dortmund, Germany, the National Human Genome Center, Howard University, Washington, D. C., U. S. A., and the Deutsche Forschungsgemeinschaft (SFB 475, “ Reduction of complexity for multivariate data structures”). The authors would like to thank Mr. Hayford-Seth Badu and Dr. Kepher Makambi for helpful suggestions.

References

[1] Bickel, P. J. and Doksum, K. A. (1977). Mathematical statistics: Basic ideas and selected topics. Holden-Day, San Francisco.

[2] Bonney, G. E. (1998). Regression analysis of disposition to correlated binary outcomes. Department of Biostatistics Technical Report No. 1998-001, Fox Chase Cancer Center, Philadelphia, PA, U.S.A.

[3] Carey, V., Zeger, S. L. and Diggle, P. (1993). Modelling multivariate binary data with alternating logistic regressions. Biometrika 80, 3, 517-526.

[4] Connolly, M. A. and Liang, K. (1988). Conditional logistic regression models for correlated binary data. Biometrika 75, 501-506.

[5] Cox, D. R. (1972). Regression Models and Life-Tables. Journal of the Royal Statistical Society, Series 34, 187-220.

[6] Cox, D. R. (1975). Partial likelihood. Biometrika 62, 2, 269-276.

[7] Garthwaite, P. H., Jolliffe, I. T. and Jones, B. (1995). Statistical inference. Prentice Hall, London, New York.

[8] Klein, J. P. and Moeschberger, M. L. (1997). Survival Analysis: Techniques for Censored and Truncated Data. Springer-Verlag, New York.

[9] Liang, K.-Y., Zeger, S. L. and Qaqish, B. (1992). Multivariate regression analysis for categorical data. Journal of the Royal Statistical Society, Series B 54, 3-40.

[10] Odai, R. N. O., Kötting, J., Kwagyan, J., Bonney, G. E. and Urfer, W. (2002). Correlated Weibull Regression Model for Multivariate Binary Data. Department of Statistics Technical Report No. 66/2002, University of Dortmund, Germany.

[11] Qaqish, B. F. and Liang, K.-Y. (1992). Marginal models for correlated binary responses with multiple classes and multiple levels of nesting. Biometrics 48, 939-950.

[12] Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. Annals of mathematical statistics, 9, 60-62.

[13] Zhao, L. P. and Prentice, R. L. (1990). Correlated binary regression using a quadratic exponential model. Biometrika 77, 642-648.