Generalized estimating equations for correlated binary data: Using the odds ratio as a measure of association

BY STUART R. LIPSITZ, NAN M. LAIRD AND DAVID P. HARRINGTON

Department of Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, U.S.A.

SUMMARY

Moment methods for analyzing repeated binary responses have been proposed by Liang & Zeger (1986), and extended by Prentice (1988). In their generalized estimating equations, both Liang & Zeger (1986) and Prentice (1988) estimate the parameters associated with the expected value of an individual's vector of binary responses as well as the correlations between pairs of binary responses. Because the odds ratio has many desirable properties, and some investigators may find the odds ratio is easier to interpret, we discuss modelling the association between binary responses at pairs of times with the odds ratio. We then modify the estimating equations of Prentice to estimate the odds ratios. In simulations, the parameter estimates for the logistic regression model for the marginal probabilities appear slightly more efficient when using the odds ratio parameterization.

Some key words: Correlated binary data; Generalized estimating equation; Marginal model; Repeated measures.

1. Introduction

Repeated measures studies arise often in biostatistical science. In these designs, subject \( i \) is observed under \( T_i \) different 'conditions' or 'times', such as at \( T_i \) occasions in a longitudinal study or with exposure to \( T_i \) drugs in a cross-over design. Note, either by design or because of missing data, \( T_i \) need not be the same for all subjects. In this paper, the response of subject \( i \) at time \( t \) is binary. The models we consider are marginal models, which relate the expected value of an individual's binary response at time \( t \), or, equivalently, the probability of success at time \( t \), to the covariates at time \( t \) via the logistic link function.

To estimate the parameters of the marginal model, Liang & Zeger (1986) and Prentice (1988) have developed moment-based approaches, which require specification of the first two moments of the vector of correlated binary responses for each individual. We review their work in § 2. In § 3, we model the correlation matrix as a function of the odds ratio for binary responses at pairs of times, and then show how to modify the estimating equations of Prentice to estimate the odds ratios. In § 4, a simulation study compares the bias and efficiency of the parameter estimates of the marginal probabilities when the correlation matrix is misspecified.
2. Review of proposed methods

First, we give notation for repeated measures data. Let \( \Omega = \{u_1, \ldots, u_T\} \) be the complete set of all \( T \) times at which any observation is made. The observations for case \( i \) \((i = 1, \ldots, N)\) are made at the subset of times \( \{a_{i1}, \ldots, a_{iT}\} \) of \( \Omega \). We define the \( T_i \times 1 \) response vector for subject \( i \) to be \( Y_i = (Y_{i1}, \ldots, Y_{iT})' \), where \( Y_{it} \) is 1 if individual \( i \) has response 1, success, at time \( a_{it} \) and is 0 if a failure. Each individual has a \( K \times 1 \) covariate vector \( x_{it} \) measured at time \( a_{it} \), which include both time-stationary and time-varying covariates. Also, we let \( X_i = (x_{i1}, \ldots, x_{iT})' \) represent the \( T_i \times K \) matrix of covariates for individual \( i \). We are interested in inferences about the parameter vector \( \beta \) associated with the vector \( \{p_i(\beta)\}' = (p_{i1}, \ldots, p_{iT}) \), which contains the marginal probabilities of success,

\[
p_{it}(\beta) = E(Y_{it}) = p_r(Y_{it} = 1 | X_i, \beta).
\]

To estimate \( \beta \), Liang & Zeger (1986) and Prentice (1988) consider generalized estimating equations of the form

\[
u(\beta) = \sum_{i=1}^{N} D_i' V_i^{-1} \{Y_i - p_{i}(\beta)\} = 0,
\]

where \( D_i = d[p_{i}(\beta)]'/d\beta \), and \( V_i \) is the ‘working’ covariance matrix of \( Y_i \). The working covariance matrix in (1) has the form

\[
V_i = A_i' R(\alpha) A_i,
\]

where

\[
A_i = \text{diag} \{\text{var} (Y_{it})\} = \text{diag} \{p_{it}(1 - p_{it})\}
\]

is specified entirely by the marginal distributions, i.e. by \( \beta \), and \( R(\alpha) = \text{corr} (Y_i) \), where \( \alpha \) represents the set of parameters associated with the model for \( \text{corr} (Y_i | \alpha) \).

To estimate \( \alpha \), note that \( \text{E}(Z_{it}) = \text{corr} (Y_{it}, Y_{it}) = \Gamma_{it}(\alpha) \), where

\[
Z_{it} = Z_{it}(\beta) = \frac{\{y_{it} - p_{it}(\beta)\} \{y_{it} - p_{it}(\beta)\}}{[p_{it}(\beta) - p_{it}(\beta)] \{1 - p_{it}(\beta)\} \{1 - p_{it}(\beta)\}].
\]

Prentice (1988) suggests a second set of estimating equations of the same form as (1) with \( Z_i = \{Z_{it}\} \) as the \( \{T_i(T_i - 1)\} \times 1 \) response vector,

\[
u(\alpha) = \sum_{i=1}^{N} C_i W_i^{-1} \{Z_i - \Gamma_i(\alpha)\} = 0,
\]

where

\[
\Gamma_i(\alpha) = \{\Gamma_{it}(\alpha)\}, \quad C_i = d\{\Gamma_i(\alpha)\}' / d\alpha
\]

and \( W_i \) is the ‘working’ covariance structure of \( Z_i \). Prentice models \( \Gamma_i(\alpha) \) as a linear function of \( \alpha \); we suggest avoiding the restrictions on the parameter space of the correlation coefficient by using the inverse of Fisher’s \( z \) transformation, namely,

\[
\Gamma_{it}(\alpha) = \{\exp (e_{it}(\alpha) - 1) / \exp (e_{it}(\alpha) + 1)\},
\]
where the $E \times 1$ vector $e_{iu}$ may be some function of $[x_{is}, x_{it}]$ and $\alpha$ is the covariate vector that affects $\Gamma_{ist}$. The estimator $[\hat{\beta}', \hat{\alpha}']$ is the solution to (1) and (3).

3. The odds ratio parameterization

An alternative to the correlation coefficient as a measure of association between pairs of binary responses is the odds ratio, which has many desirable properties (Bishop, Fienberg & Holland, 1975, Ch. 11) and possibly is easier to interpret than the correlation coefficient. First, we let the set of estimating equations (1) for the marginal parameters $\beta$ remain the same. Now, suppose instead of having $Z_i = \{Z_{ist}\}$ as the response vector for the second set of estimating equations, we have the $\{T_i(T_i - 1)\} \times 1$ random vector $U_i = \{Y_{ist}\}$, where

$$Y_{ist} = I\{Y_{ia} = 1, Y_{it} = 1\} = Y_{is} Y_{it}$$

and $I\{\}$ is an indicator function. To use a second set of estimating equations similar to (3), we must specify $E(U_i)$ and $W_i = \text{var}(U_i)$. Now,

$$p_{ist} = E(Y_{ist}) = \text{pr}\{Y_{ia} = 1, Y_{it} = 1\}$$

is the joint probability of success at both times $s$ and $t$; given a model for $p_{ist}$, along with the ‘marginal model’ for $p_{ia}$ and $p_{it}$, the $(s, t)$th element of $R_i(\alpha)$ in (1) is

$$\Gamma_{ist} = \text{corr}(Y_{is}, Y_{it}) = \frac{p_{ist} - p_{ia} p_{it}}{\{p_{ia}(1 - p_{ia}) p_{it}(1 - p_{it})\}^{1/2}}.$$  (4)

Now, we show how to write $p_{ist}$ as a function of $p_{ia}$, $p_{it}$ and the odds ratio between the responses at times $s$ and $t$. We can form a $2 \times 2$ table by cross-classifying an individual’s responses at times $s$ and $t$; the odds ratio for the $2 \times 2$ table is

$$\tau_{ist} = \frac{p_{ist}(1 - p_{ia} - p_{it} + p_{ist})}{(p_{is} - p_{ist})(p_{it} - p_{ist})}.$$  (5)

Using the quadratic formula, we can solve for $p_{ist}$ in terms of the odds ratio, $\tau_{ist}$, and the two marginal probabilities, $p_{ia}$ and $p_{it}$. The solution (Mardia, 1967) that will always lie in $[0, 1]$ is

$$p_{ist} = \begin{cases} f_{ist} - \frac{\{\tau_{ist} - 1\} (p_{ia} - p_{it})}{2(\tau_{ist} - 1)} & (\tau_{ist} \neq 1), \\ \tau_{ist} p_{ia} p_{it} & (\tau_{ist} = 1), \end{cases}$$  (6)

where $f_{ist} = \{(1 - (1 - \tau_{ist})(p_{ia} + p_{it})\}$. Note that $p_{ist} = p_{ist}(\beta, \alpha)$ is a function of $\beta$ through $p_a(\beta)$ and $p_t(\beta)$, and of $\alpha$ through the odds ratio $\tau_{ist}(\alpha)$, where $\alpha$ can now be thought of as the parameters associated with the odds ratios. One can model log $\tau_{ist}$ as a linear function of $\alpha$ and the covariate vector $e_{ist}$.

Given the model for $p_{ist} = E(Y_{ist})$, we must determine how to specify $W_i = \text{var}(U_i)$. Since $Y_{ist}$ is a binary random variable, its variance is $\text{var}(Y_{ist}) = p_{ist}(1 - p_{ist})$. Consider $\text{cov}(Y_{ist}, Y_{iuv})$ for four different times $s, t, u$ and $v$. It is easy to show that

$$\text{cov}(Y_{ist}, Y_{iuv}) = \text{pr}\{Y_{ia} = 1, Y_{it} = 1, Y_{iu} = 1, Y_{iv} = 1\} - p_{ist} p_{iuv}.$$
Yet, this covariance is a function of a fourth order probability, which is a nuisance parameter in the covariance matrix of $U_t$ which we would rather not estimate. Thus, we suggest specifying

$$W_t = \text{diag} \{ \text{var} (Y_{it}) \} = \text{diag} \{ p_{it}(1 - p_{it}) \}.$$ 

Then our second set of estimating equations is

$$u(\alpha) = \sum_{i=1}^{N} C_i W_t^{-1} \{ U_i - \theta_i(\beta, \alpha) \} = 0,$$  
(7)

where $\theta_i(\beta, \alpha)$ is the model for $E(U_i)$ and $C_i = d\theta_i / d\alpha$. Although $\theta_i$ is a function of both $\alpha$ and $\beta$, we assume $\beta$ is fixed in $\theta_i$ in (7) and, in $C_i$, we have only taken derivatives of $\theta_i$ with respect to $\alpha$. In Prentice's second set of estimating equations (3), $\hat{\beta}$ is assumed fixed in $Z_{it}$. This 'fixing' of $\beta$ in a second set (3) or (7) does not affect the asymptotics discussed below, which require that $p_i(\beta)$ is correctly specified and $\hat{\alpha}$ is consistent for some vector which we have denoted by $\alpha$. A Fisher-scoring type algorithm to compute $[\hat{\beta}', \hat{\alpha}']$ can be used

$$\hat{\beta}^{(m+1)} = \hat{\beta}^{(m)} - \left( \sum_{i=1}^{N} \hat{C}_i W_t^{-1} \right)^{-1} \left[ \sum_{i=1}^{N} \hat{C}_i W_t^{-1} \{ Y_i - p_i(\hat{\beta}) \} \right],$$  
(8)

$$\hat{\alpha}^{(m+1)} = \hat{\alpha}^{(m)} - \left( \sum_{i=1}^{N} \hat{C}_i W_t^{-1} \right)^{-1} \left[ \sum_{i=1}^{N} \hat{C}_i W_t^{-1} \{ U_i - \theta_i(\hat{\beta}, \hat{\alpha}) \} \right].$$  
(9)

Using expansions similar to Prentice (1988), assuming that the regression for $Y_i$ is correctly specified, $[\hat{\beta}', \hat{\alpha}']$ is consistent for $[\beta', \alpha']$, and also $N^1[(\beta - \beta)', (\alpha - \alpha)']$ has an asymptotic distribution which is multivariate normal with mean vector 0 and covariance matrix

$$V_{\beta, \alpha} = \lim_{N \to \infty} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{21} \\ B_{12} & B_{22} \end{bmatrix},$$  
(10)

where

$$B_{11} = N^{-1} \sum_{i=1}^{N} D_i' V_i^{-1} D_i, \quad B_{22} = N^{-1} \sum_{i=1}^{N} C_i W_i^{-1} C_i,$n

$$B_{21} = B_{22}^{-1} \left\{ N^{-1} \sum_{i=1}^{N} C_i W_i^{-1} (-d\theta_i / d\beta) \right\} B_{11}^{-1},$$

$$\Sigma_{11} = N^{-1} \sum_{i=1}^{N} D_i' V_i^{-1} \{ \text{var} (Y_i) \} V_i^{-1} D_i, \quad \Sigma_{22} = N^{-1} \sum_{i=1}^{N} C_i W_i^{-1} \{ \text{var} (U_i) \} W_i^{-1} C_i,$n

$$\Sigma_{12} = N^{-1} \sum_{i=1}^{N} D_i' V_i^{-1} \{ \text{cov} (Y_i, U_i) \} W_i^{-1} C_i.$$

The asymptotic covariance matrix of $N^1(\hat{\beta} - \beta)$ is

$$V_{\beta} = \lim_{N \to \infty} (B_{11}^{-1} \Sigma_{11} B_{11}^{-1}),$$  
(11)

and if $\text{var} (Y_i)$ is correctly specified, $\Sigma_{11} = B_{11}$, and (11) reduces to

$$V_{\beta} = \lim_{N \to \infty} B_{11}^{-1}.$$  
(12)
Now, \( V_{p,a} \) can be consistently estimated by replacing \( \beta \) and \( \alpha \) by their estimates, and also \( \text{var}(Y_i) \) by \( (Y_i - \hat{p}_i)(Y_i - \hat{p}_i)' \), \( \text{var}(U_i) \) by \( (U_i - \theta_i)(U_i - \theta_i)' \) and \( \text{cov}(Y_i, U_i) \) by \( (Y_i - \hat{p}_i)(U_i - \hat{\theta}_i)' \).

Finally, examples for both the correlation and the odds ratio parameterizations for \( p_{\alpha} \) can be constructed in which \( V_i(\beta, \alpha) \) in (1) is not positive definite for all possible values of \( (\beta, \alpha) \); thus \( (\beta, \alpha) \) is not a product space. However, in most practical situations, the solution \( (\hat{\beta}, \hat{\alpha}) \) will be such that \( V_i(\hat{\beta}, \hat{\alpha}) \) is positive definite.

### 4. Simulation study

We used simulation to examine the effect on estimating \( \beta \) of the misspecification of the association structure \( \alpha \). For ease in specifying the underlying true model, we considered only paired binary data, that is \( T = 2 \), and \( j = 1, \ldots, 4 \) covariate strata. In each replication in the simulations, we drew a sample of size 25 from each of the 4 strata. The probability model for the marginal probabilities in stratum \( j \) was logit \( (p_{it}) = \beta x_{ij} \), where \( \beta = 1 \) and thus \( x_{ij} = \text{logit}(p_{ij}) \); we let

\[
(p_{11}, p_{12}) = (0.1, 0.1), \quad (p_{21}, p_{22}) = (0.5, 0.5), \quad (p_{31}, p_{32}) = (0.2, 0.4),
\]

\[
(p_{41}, p_{42}) = (0.5, 0.3).
\]

We performed seven sets of simulations with 1000 replicates each; three sets in which the true model had a constant odds ratio over all strata, with a value of 2, 5 or 10; three sets in which the true model had a constant correlation coefficient over all strata, with a value of 0.1, 0.3 or 0.6; and one set under independence of \( Y_1 \) and \( Y_2 \). Results are given in Table 1.

For each of the seven true probability models, we estimated \( \beta \) under three ‘working models’; independence, constant correlation coefficient across strata, and constant odds ratios across strata. We used estimating equations (1) and (7) to fit the models, specifying \( W_i = \text{diag}\{p_{it}(1 - p_{it})\} \) in (7). Also in (7), we specified \( p_{112} \) as in (6) for the odds ratio model, and \( p_{112} \) as

\[
p_{112} = p_{11}p_{12} + \Gamma\{p_{11}(1 - p_{11})p_{12}(1 - p_{12})\}, \quad (13)
\]

for the correlation coefficient model.

We then tabulated the bias in \( \hat{\beta} \) from the three ‘working’ models. We also examined nominal 95% confidence intervals for \( \beta \), using both (11) and (12), and \( t_{0.025,99} = 1.9842 \), the 97.5th percentile of the \( t \) distribution with \( N - 1 = 99 \) degrees of freedom. We let \( I_b \) be 1 if the confidence interval for the \( b \)th simulation contains \( \beta \), and 0 otherwise, and looked at the average of the \( I_b \) over the simulations. With 1000 replications, using large sample normal theory, the average of the \( I_b \)'s should be in [93.5, 96.4] 95% of the time if the coverage probability is 95%.

Table 1 shows that all the estimators of \( \beta \) are biased slightly upward, although this bias probably is of little practical importance. As the true strength of association gets bigger, the estimators of \( \beta \) tend to be more biased. The bias of three ‘working estimators’ appears comparable, and thus there appears to be no benefit to adopting the estimator on the basis of bias. Further, the working estimator assuming a common odds ratio, \( \hat{\beta}_{OR} \), appears to have slightly smaller variance than the working estimator assuming a common correlation coefficient, \( \hat{\beta}_C \), for most of the true association parameterizations; however, these sample variances are similar, and the differences could be due to sampling variability.
Table 1. Simulation results of general estimating equation estimator for logit \( p_{ij} = x_{ij}\beta, \beta = 1, \) for strata \( j = 1, \ldots, 4, \) each with sample size 25, and \( t = 1, 2 \)

<table>
<thead>
<tr>
<th>True association model</th>
<th>Working association model</th>
<th>Constant odds ratio</th>
<th>Constant correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independence</td>
<td>Independence</td>
<td>1.020 (0.0310)</td>
<td>1.023 (0.0312)</td>
</tr>
<tr>
<td></td>
<td>96.1</td>
<td>96.1</td>
<td>95.9</td>
</tr>
<tr>
<td></td>
<td>95.1</td>
<td>95.2</td>
<td>95.3</td>
</tr>
<tr>
<td>( \tau = 2 )</td>
<td>1.019 (0.0357)</td>
<td>1.021 (0.0357)</td>
<td>1.020 (0.0357)</td>
</tr>
<tr>
<td></td>
<td>95.0</td>
<td>96.2</td>
<td>96.3</td>
</tr>
<tr>
<td></td>
<td>95.4</td>
<td>95.7</td>
<td>95.6</td>
</tr>
<tr>
<td>( \tau = 5 )</td>
<td>1.022 (0.0397)</td>
<td>1.024 (0.0382)</td>
<td>1.025 (0.0391)</td>
</tr>
<tr>
<td></td>
<td>93.3</td>
<td>95.5</td>
<td>96.0</td>
</tr>
<tr>
<td></td>
<td>95.1</td>
<td>95.1</td>
<td>95.3</td>
</tr>
<tr>
<td>( \tau = 10 )</td>
<td>1.029 (0.0428)</td>
<td>1.027 (0.0403)</td>
<td>1.031 (0.0427)</td>
</tr>
<tr>
<td></td>
<td>92.7</td>
<td>95.1</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>94.1</td>
<td>94.2</td>
<td>94.8</td>
</tr>
<tr>
<td>( R = 0.1 )</td>
<td>1.015 (0.0365)</td>
<td>1.018 (0.0365)</td>
<td>1.017 (0.0362)</td>
</tr>
<tr>
<td></td>
<td>94.7</td>
<td>95.3</td>
<td>95.5</td>
</tr>
<tr>
<td></td>
<td>94.9</td>
<td>94.9</td>
<td>94.9</td>
</tr>
<tr>
<td>( R = 0.3 )</td>
<td>1.027 (0.0401)</td>
<td>1.028 (0.0396)</td>
<td>1.028 (0.0401)</td>
</tr>
<tr>
<td></td>
<td>93.8</td>
<td>95.3</td>
<td>95.8</td>
</tr>
<tr>
<td></td>
<td>95.0</td>
<td>95.5</td>
<td>95.6</td>
</tr>
<tr>
<td>( R = 0.6 )</td>
<td>1.038 (0.0498)</td>
<td>1.032 (0.0433)</td>
<td>1.031 (0.0436)</td>
</tr>
<tr>
<td></td>
<td>89.8</td>
<td>95.3</td>
<td>94.2</td>
</tr>
<tr>
<td></td>
<td>95.4</td>
<td>94.5</td>
<td>95.0</td>
</tr>
</tbody>
</table>

First row in each cell contains sample mean of estimates of \( \beta \) over the simulations, and sample variance in parentheses. Second row contains proportion of 95% confidence intervals containing true value when using (12) to estimate variance. Third row contains proportion of 95% confidence intervals containing true value when using robust variance estimate (11).

Alternatively, because the odds ratio, and not the correlation coefficient, is orthogonal to the marginal parameters when \( T = 2 \) (McCullagh & Nelder, 1989, p. 228) we might expect the estimate \( \hat{\beta}_{OR} \) to vary less as the odds ratio varies than we would expect \( \hat{\beta}_C \) to vary as the correlation coefficient varies. This could account for \( \hat{\beta}_{OR} \) having slightly less variability than \( \hat{\beta}_C \) in the simulations.

The estimator \( \hat{\beta}_I \) under the working independence assumption is slightly less biased than \( \hat{\beta}_{OR} \) and \( \hat{\beta}_C \) except when the true association is strong, i.e. the odds ratio is 10 or the correlation coefficient is 0.6. Also, \( \hat{\beta}_I \) appears to have similar variability to \( \hat{\beta}_{OR} \) or \( \hat{\beta}_C \) when the association is weak, but appears to be only 87% as efficient as \( \hat{\beta}_{OR} \) when the true common correlation coefficient is 0.6, as measured by the ratio of the same sample variances.

Although we cannot make strict recommendations based on these simulations, we suggest the following. We attempted to perform simulations with \( N \) and \( T \) both small, that is \( T = 2, N = 40, \) with 10 in each stratum, but found that (8) and (9) did not converge for either the working odds ratio or working correlation model for all replications, although the working independence model did converge for all of the replications. With
small $N$ and $T$, zero cells are much more likely, and $\alpha$ may be highly variable and/or inestimable because of boundary value problems. Also, zero cells are more likely when the association between the responses at the two times is strong, in which both $\Pr \{ Y_{11} = 1, Y_{12} = 0 \}$ and $\Pr \{ Y_{11} = 0, Y_{12} = 1 \}$ are small. In the simulations with $N = 40$, as the association between the responses at the two times increased, (8) and (9) converged less often. For example, with a common correlation coefficient of 0.6, the estimated log odds ratio tended to $\infty$ in about 5% of the replications and the estimated correlation coefficient to 1 in about 4% of the replications; when the common correlation coefficient was 0.1, only about 1% of the estimated log odds ratios or correlation coefficients converged to boundary points. These problems occurred when most, or all, $Y_{11}$ and $Y_{12}$ were equal in a replication. Thus, when $N$ and $T$ are both small and one is only interested in estimating $\beta$, one should probably use the working assumption of independence. Exploration into the small sample properties of the solutions to (1) and (7) is a topic for future research.

When $N$ is of sufficient size and much larger than $T$ as in our simulations, the coverage probabilities using the robust estimator of variance (11) appears correct in all cases, and the variance estimator given by (12) appears correct when it is appropriate. However (Prentice, 1988), when $T$ is greater than or equal to $N$, the robust estimator (12) will not be very accurate, and one must rely on (11), so that the odds ratio may provide a viable alternative to the correlation coefficient.

Finally, note, with $T = 2$, $[\beta, \alpha]$ is a product space for the odds ratio model; $[\beta, \alpha]$ for the correlation model in (13) is not a product space (Kupper & Haseman, 1979). However, we found that $V_i(\hat{\beta}, \hat{\alpha})$ was positive definite for every replication of every simulation, which supports our claim at the end of § 3 that the restrictions on the parameter space is usually of little practical importance.

5. DISCUSSION

The methods that we propose are useful in that they allow the estimation of pairwise odds ratios in the analysis of correlated binary data. When only the marginal parameters $\beta$ are of interest, our simulations show that $\hat{\beta}_{\text{OR}}$ is slightly more efficient than $\hat{\beta}_{\text{C}}$. However, often the pairwise correlation coefficients or odds ratio are of interest themselves. Further, regressions for the pairwise association parameters with covariate vector $\{e_{it}\}$ may be of interest. We have shown how to develop equations similar to those of Prentice to estimate the parameters $\alpha$ of the pairwise log odds ratios. We found that the computing required for either our proposed odds ratio parameterization or a correlation parameterization was about the same. Thus, we feel that one parameterization should not be chosen over the other because of computational issues; we feel one should choose the association parameterization for purposes of interpretation. Finally, the methods discussed in § 3 allow any measure of association between pairs of binary random variables in the parameterization of $p_{it}$ (Bishop et al., 1975, pp. 14-6), such as the relative risk if the $T$, conditions have an underlying order as in a longitudinal study.

ACKNOWLEDGEMENTS

This work was supported in part by the United States National Institutes of Health. The authors would like to thank the referees and editor for their helpful suggestions.
REFERENCES


[Received May 1989. Revised July 1990]