Conditional Inference Functions for Mixed-Effects Models with Unspecified Random-Effects Distribution

Peng Wang
Bowling Green State University
Joint Work with Annie Qu and Cindy Tsai

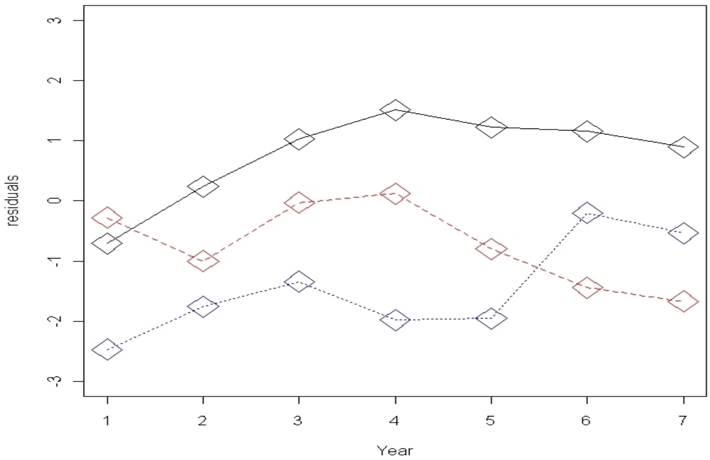
Current Challenges in Statistical Learning

Banff International Research Station
for Mathematical Innovation and Discovery

Motivating Example

- A longitudinal observational study, non-surgical periodontal treatment effect on tooth loss
- There are 722 subjects for 7-year follow up
- The main covariate: non-surgical periodontal treatment (1 or 0) for three years before the study
- Other covariates:
  - Gender
  - Age
  - Variables to measure teeth health condition
- There is subject-specific variation among subjects

A Graph of Longitudinal Data



Longitudinal Data

- Tooth loss and other covariates are recorded repeatedly over a 7-year period
- Measurements within the same subject are correlated
- Major approaches for correlated data:
  - Marginal models
  - Mixed-effects models

Marginal Models

- The inference of the population average is the main focus
- Generalized Estimating Equations (GEE) (Liang & Zeger, 1986); Quadratic Inference Functions (Qu et al., 2000):
  - Does not require likelihood function
  - Consistent even if the correlation structure is misspecified
  - Estimator is efficient with the correct working correlation
  - Provides robust sandwich variance estimator

Mixed Models

- There is heterogeneity among subjects
- Able to incorporate several sources of variation: random effects and serial correlation
- Limitations:
  - Requires parametric assumption for random effects, usually normality assumption
  - Involves high dimensional integration for non-normal random effects

Existing Methods for Generalized Linear Mixed-Effects Model

- Penalized quasilikelihood (PQL) (Breslow and Clayton, 1993)
- Hierarchical generalized linear model (HGLM) (Lee and Nelder, 1996, 2001)
- Conditional likelihood (Jiang, 1999)
- Conditional second-order generalized estimating equations (Vonesh et al., 2002)

Limitations and Assumptions

- Require normal assumption for random effects (PQL, second order GEE).
- Require estimation of variance components (PQL and conditional second order GEE).
- Do not incorporate serial correlation (PQL, HGLM and conditional likelihood).

Advantages of the Proposed Approach

- A new approach using the conditional quadratic inference function
- Does not require distribution assumption of random effects
- Does not require the likelihood function, only involves the first two moments
- Accommodates variations from both random effects and serial correlations
- Does not require estimation of unknown variance components or correlation parameters
- Challenge: the dimension of random effects parameters increases as the sample size increases

GEE

Generalized estimating equations (Liang & Zeger,1986) can be represented as

$$\sum_{i=1}^{N} \left(\frac{\partial \mu_i}{\partial \beta}\right)' A_i^{-1/2} R^{-1}(\alpha) A_i^{-1/2} (y_i - \mu_i) = 0,$$

where $y_i = (y_{i1}, ..., y_{it})$ is the response vector for the ith subject,
$\mu_i = E(y_i) = (\mu_{i1}, \ldots, \mu_{it})$ is the mean vector for the ith subject,
A_i is a diagonal matrix of variance components of y_i,
and $R(\alpha)$ is the working correlation

Representation of Correlation Matrix

- Approximate R^{-1} by $\sum_{j=1}^{m} a_j M_j$
  - $M_1, \ldots, M_m$ are known basis matrices
  - $a_1, \ldots, a_m$ are unknown constants
- The linear representation can accommodate most common working correlation structures such as AR-1, exchangeable or block diagonal

QIF Approach (Qu et al., 2000)

- GEE: $\sum_{i=1}^{N} \left(\frac{\partial \mu_i}{\partial \beta}\right)' A_i^{-1/2} R^{-1}(\alpha) A_i^{-1/2}(y_i - \mu_i) = 0$
- Substitute $R^{-1} \approx \sum_{j=1}^{m} a_j M_j$ into GEE,

$$g = \sum \dot{\mu_i}' A_i^{-1/2} (\sum_{j=1}^{m} a_j M_j) A_i^{-1/2} (y_i - \mu_i)$$

QIF Approach

- Define the extended score

$$\bar{G}_N(\beta) = \frac{1}{N}\sum g_i(\beta) = \frac{1}{N}\begin{pmatrix} \sum(\dot{\mu}_i)' A_i^{-1/2} M_1 A_i^{-1/2}(y_i - \mu_i) \\ \vdots \\ \sum(\dot{\mu}_i)' A_i^{-1/2} M_m A_i^{-1/2}(y_i - \mu_i) \end{pmatrix}$$

- The GEE is a linear combination of $\bar{G}_N(\beta)$
- The QIF estimator $\hat{\beta} = \arg\min \bar{G}_N' C_N^{-1} \bar{G}_N$, where $C_N = (1/N)\sum g_i(\beta) g_i'(\beta)$
- The QIF estimator $\hat{\beta}$ is more efficient than the GEE estimator under the misspecified correlation structure
- It provides an objective and inference function for model checking and testing

Mixed-Effects Model

A mixed effects model conditional on random effects b_i for longitudinal data is modeled as

$$E(y_{it}|x_{it}, b_i) = \mu(x_{it}'\beta + z_{it}'b_i), i = 1, ...N, t = 1, ..., n_i$$

- y_{it} is the response variable
- x_{it} are the covariates
- z_{it} are the covariates for random effects
- β are the fixed-effect parameters
- $b = (b_1, .., b_N)$ are the random-effects parameters, have the same order of dimension as the sample size

Penalized Conditional Quasilikelihood

- The conditional quasi-likelihood of y given the random effects b is $l_q^b = -\frac{1}{2\phi}\sum_{i=1}^{N} d_i(y_i, \mu_i^b)$, where $d_i(y,u) = -2\int_y^u \frac{y-u}{a_i v(u)} du$
- Require a constraint to ensure identifiability: $P_A b = 0$
- P_A is the projection matrix on the null space of $(I - P_X)Z$

Penalized conditional quasilikelihood (Jiang, 1999)

$$l_q = -\frac{1}{2\phi}\sum_{i=1}^{N} d_i(y_i, \mu_i^b) - \frac{1}{2}\lambda|P_A b|^2$$

- The penalty λ is fixed, and is chosen as 1 in Jiang (1999)
- Jiang's approach does not converge

Conditional Extended Score Corresponding for β and b

- Take the derivatives of the penalized conditional quasilikelihood l_q corresponding to β and b
- The quasi-score equation corresponding to the fixed effect β is

$$\sum_{i=1}^{N}(\frac{\partial \mu_i^b}{\partial \beta})'(\mathrm{W}_i^b)^{-1}(y_i - \mu_i^b) = 0.$$

- The quasi-score equation corresponding to the random effects b is

$$\left(\begin{array}{ccc} h_1 & = & (\frac{\partial \mu_1^{b_1}}{\partial b_1})'(\mathrm{W}_1^b)^{-1}(y_1 - \mu_1^{b_1}) - \lambda \frac{\partial P_A b}{\partial b_1} P_A b = 0 \\ & & \vdots \\ h_N & = & (\frac{\partial \mu_i^{b_N}}{\partial b_N})'(\mathrm{W}_N^b)^{-1}(y_N - \mu_N^{b_N}) - \lambda \frac{\partial P_A b}{\partial b_N} P_A b = 0 \end{array}\right)$$

Extended Score for β

- Construct extended scores associated with the fixed effect β

$$G_N^f = \frac{1}{N}\sum_{i=1}^{N} g_i^f(\beta) = \frac{1}{N}\begin{pmatrix} \sum_{\mathrm{i=1}}^{\mathrm{N}}(\frac{\partial\mu_{\mathrm{i}}^{\mathrm{b}}}{\partial\beta})'\mathrm{A}_{\mathrm{i}}^{-1/2}\mathrm{M}_1\mathrm{A}_{\mathrm{i}}^{-1/2}\left(\mathrm{y_i}-\mu_{\mathrm{i}}^{\mathrm{b}}\right) \\ \vdots \\ \sum_{\mathrm{i=1}}^{\mathrm{N}}(\frac{\partial\mu_{\mathrm{i}}^{\mathrm{b}}}{\partial\beta})'\mathrm{A}_{\mathrm{i}}^{-1/2}\mathrm{M}_{\mathrm{m}}\mathrm{A}_{\mathrm{i}}^{-1/2}\left(\mathrm{y_i}-\mu_{\mathrm{i}}^{\mathrm{b}}\right) \end{pmatrix}$$

- Conditional on b,

$$\hat{\beta} = \arg\min(\bar{G}_N^f)'(\bar{C}_N^f)^{-1}(\bar{G}_N^f)$$

where $\bar{C}_N^f = (1/N)\sum g_i^f(\beta)g_i^f(\beta)'$

Extended Score Corresponding to b

- For the ith subject, the quasi-score associated with the random effect:

$$h_i = (\frac{\partial \mu_i^{b_i}}{\partial b_i})'(\mathrm{W}_1^b)^{-1}(y_1 - \mu_1^{b_i}) - \lambda \frac{\partial P_A b}{\partial b_i} P_A b = 0$$

- Substitute $W_i = A_i^{\frac{1}{2}} R A_i^{\frac{1}{2}}$ and assume independent structure for R
- The extended score for the random effect b for subject i

$$g_i^r = \begin{pmatrix} (\frac{\partial \mu_i^{b_i}}{\partial b_i})' A_i^{-1}(y_i - \mu_i^{b_i}) \\ \lambda \frac{\partial P_A b}{\partial b_i} P_A b \end{pmatrix}$$

- In a simple random intercept model, $\frac{\partial P_A b}{\partial b_i} P_A b = \sum_{i=1}^N b_i / N$
- Jiang (1999) only considers the constraint for the random effect $P_A b = 0$
- This constraint is not sufficient to ensure algorithm convergence

Extended Score for b

- The convergence problem becomes more serious when there are high-dimensional random effects involved in the model
- We include an addditional penalty term λb_i which also controls the variance of the random effects estimators to ensure that the algorithm converges
- The new extended scores for b are

$$g^r = \left\{(g_1^r)', \lambda b_1', \ldots, (g_N^r)', \lambda b_N'\right\}'$$

- For given fixed effects β, $\hat{b} = \arg\min(g^r)'(g^r)$
- No replicate for each g_i^r, so there is no weighting matrix in the estimation

Regularity Conditions

- The parameter space S is compact
- There is a unique $\beta_0 \in S$ which satisfies $E[g(\beta_0|b_0)] = 0$
- The derivative of the score function $\dot{g}_{i,b}(\hat{\beta}|b_0) = O_p(1)$
- Expectation of the continuous score $E[g(\beta|b)]$ is continuous and differentiable in both β and b
- The weighting matrix $C_N(\beta|b) \rightarrow_{a.s.} C_0(\beta|b)$ and $A_N(\beta|b) \rightarrow_{a.s.} A_0(\beta|b)$, where $C_0^{-1}(\beta|b) = A_0(\beta|b)A_0(\beta|b)'$
- The estimating functions conditional on the estimated random effects converges to 0 in probability

$$E[E\{g_i(\beta_0|\hat{b})\}] \xrightarrow{p} 0 \quad \text{as } N \rightarrow \infty$$

- This condition is much weaker than the consistency for the random effects estimator

Asymptotic Properties

Theorem 1: Under some regularity conditions, the QIF estimator for the fixed effects $\hat{\beta}_1$ has the following properties as $N \to \infty$

I. (Consistency) $\hat{\beta}_1 \to_p \beta_0$.

II. (Asymptotic Normality) $\sqrt{N}(\hat{\beta}_1 - \beta_0) \xrightarrow{d} N(0, \Omega_1)$

Difficulties:

- No normality assumption for the random effects
- $\hat{b}$ is not required to be a consistent estimator for true b

III If $\hat{b}$ is a consistent estimator of b_0, then $\Omega_1 = \lim_{n,N\to\infty} \ddot{Q}^{-1}_{\beta\beta}(\hat{\beta}_1|\hat{b}) = \Omega_0$, where $\ddot{Q}^{-1}_{\beta\beta}(\hat{\beta}_1|\hat{b}) \approx \{\dot{G}_{N,\beta}(\hat{\beta}_1|\hat{b}) C_N^{-1}(\hat{b}) \dot{G}_{N,\beta}(\hat{\beta}_1|\hat{b})\}^{-1}$

Algorithm

- For a fixed λ, the iterative algorithms are
  1. Start with an initial estimator $\hat{\beta}$
  2. Given $\hat{\beta}$ as the fixed effects estimator, estimate the random effects by minimizing $(g^r)'(g^r)$, obtain random effects estimator $\hat{b}$
  3. Given random effect estimator $\hat{b}$ obtained in Step 2, update $\hat{\beta}$ by minimizing $(\bar{G}_N^f)'(\bar{C}_N^f)^{-1}(\bar{G}_N^f)$, iterate between Step 2 and 3 until convergence is reached
- The tuning parameter λ can be chosen by minimizing a BIC-type of criteria

$$N(\bar{G}_N^f)'(\bar{C}_N^f)^{-1}(\bar{G}_N^f) + (\log N)(P_A b)'\Sigma_b^{-1}(P_A b).$$

Handling Unbalanced Data

- For unbalanced longitudinal data, treat as missing data case
- Create the largest cluster with a size T which contains time points for all possible measurements
- Let T_i be the ith cluster size, define the $T \times T_i$ transformation matrix Λ_i
- By removing the columns of the identity matrix corresponding to the missing observations

Handling Unbalanced Data

- Define $y_i^* = \Lambda_i y_i$, $\mu_i^*(\tilde{\beta}) = \Lambda_i \mu_i(\tilde{\beta})$, and $\dot{\mu}_i^*(\tilde{\beta}) = \Lambda_i \dot{\mu}_i(\tilde{\beta})$
- Components in y_i^*, μ_i^* and $\dot{\mu}_i^*$ are the same as for the non-missing observations but are 0 for the missing components
- Let $A_i^* = \Lambda_i A_i \Lambda_i^T$ and $(A_i^*)^{-1} = \Lambda_i A_i^{-1} \Lambda_i^T$, the marginal variance A_i^* is 0 for the missing observations
- Assume that R is the common working correlation matrix for the fully observed responses
- The basis matrices $M_1, \ldots, M_m$ could still be used to model the inverse of the correlation matrix for all clusters
- The utilization for these basis matrices are different for different clusters
- After transformation, the cluster size of the unbalanced data becomes equal

Simulation Set-up for Binary Data

- Sample size: 100 ($i = 1, \ldots, 100$); cluster size: 4 or 5
- The conditional correlated binary outcomes are generated from

$$\text{logit}(\mu_i^b) = \beta_0 + b_{0i} + x_i(\beta_1 + b_{1i}),\ \texttt{corr}(y_i|x_i, b_{0i}, b_{1i}) = R$$

- The covariate x_i is generated from a uniform (0.5, 1.5), $\beta_0 = -0.3$, $\beta_1 = 0.3$
- Correlation structures: independent, exchangeable, or AR-1, correlation parameter $\rho = 0.7$
- Both random intercept b_{0i} and random slope b_{1i} are from a bimodal distribution of a rescaled Beta(0.5, 0.5)
- Apply mixed QIF with three types of working correlations, PQL, the SAS GLIMMIX and the NLMIXED

Results for Fixed-Effects β_0

Table: MSE for the estimator of the intercept $\beta_0 = -0.3$ for binary responses when $\rho = 0.7$ from 200 simulations.

Method	$N = 100$ True correlation		
	Independent	Exchangeable	AR-1
QIF (ind)	0.1464	0.1373	0.1298
QIF (exch)	0.1494	0.0762	0.1080
QIF (AR-1)	0.1499	0.0842	0.0802
PQL	0.1517	1.8556	0.6628
GLIMMIX (Ind)	0.1604[1]	1.0159	0.3912[2]
GLIMMIX (AR-1)	0.1713[3]	1.0616[4]	0.1226[5]
NLMIXED	0.1505[6]	1.1474	0.5126

Number of non-convergence outcomes from GLIMMIX and NLMIXED procedures are tabulated as follows: **1.** 173; **2.** 7; **3.** 174; **4.** 1; **5.** 174; **6.** 7.

Results for Fixed-Effects β_1

Table: MSE for the estimator of the intercept $\beta_1 = 0.3$ for binary responses when $\rho = 0.7$ from 200 simulations.

Method	$N = 100$ True correlation		
	Independent	Exchangeable	AR-1
QIF (ind)	0.1414	0.1072	0.0979
QIF (exch)	0.1425	0.0534	0.0732
QIF (AR-1)	0.1447	0.0578	0.0494
PQL	0.1451	1.1949	0.4354
GLIMMIX (Ind)	0.1445[1]	1.2468	0.3509[2]
GLIMMIX (AR-1)	0.1656[3]	1.2860[4]	0.0755[5]
NLMIXED	0.1448[6]	0.9600	0.3971

Number of non-convergence outcomes from GLIMMIX and NLMIXED procedures are tabulated as follows: **1.** 173; **2.** 7; **3.** 174; **4.** 1; **5.** 174; **6.** 7.

Results for Random-Effects of β_0

Table: Mean and the standard errors of the variance component estimator of β_0 for binary responses when $\rho = 0.7$ from 200 simulations. The true variance of the random intercept is 0.015

Method	$N = 100$ True correlation		
	Independent	Exchangeable	AR-1
QIF (ind)	$0.0051_{0.0000}$	$0.0153_{0.0000}$	$0.0133_{0.0000}$
QIF (exch)	$0.0051_{0.0000}$	$0.0154_{0.0000}$	$0.0133_{0.0000}$
QIF (AR-1)	$0.0051_{0.0000}$	$0.0154_{0.0000}$	$0.0134_{0.0000}$
PQL	$0.4602_{0.0706}$	$52.6074_{5.8626}$	$12.4455_{1.0486}$
GLIMMIX (Ind)	$0.0812^{\mathbf{1}}_{0.0043}$	$11.2695_{0.2750}$	$4.8532^{\mathbf{2}}_{0.0979}$
GLIMMIX (AR-1)	$0.1622^{\mathbf{3}}_{0.0084}$	$11.4377^{\mathbf{4}}_{0.3306}$	$0.5640^{\mathbf{5}}_{0.0313}$
NLMIXED	$0.0515^{\mathbf{6}}_{0.0067}$	$25.4358_{0.6391}$	$8.5088_{0.2236}$

Number of non-convergence outcomes from GLIMMIX procedures are tabulated as follows: **1.** 173; **2.** 7; **3.** 174; **4.** 1; **5.** 174; **6.** 7.

Results for Random-Effects of β_1

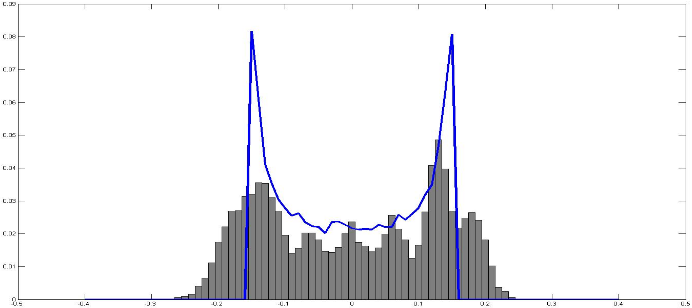
Table: Mean and the standard errors of the variance component estimator of β_1 for binary responses when $\rho = 0.7$ from 200 simulations. The true variance of the random intercept is 0.015

Method	$N = 100$ True correlation		
	Independent	Exchangeable	AR-1
QIF (ind)	$0.0073_{0.0000}$	$0.0262_{0.0000}$	$0.0212_{0.0000}$
QIF (exch)	$0.0073_{0.0000}$	$0.0263_{0.0000}$	$0.0213_{0.0000}$
QIF (AR-1)	$0.0073_{0.0000}$	$0.0263_{0.0000}$	$0.0213_{0.0000}$
PQL	$0.3573_{0.0592}$	$23.6498_{2.0657}$	$3.0542_{0.3737}$
GLIMMIX (Ind)	$0.0763^{\mathbf{1}}_{0.0049}$	$8.9704_{0.3339}$	$2.6071^{\mathbf{2}}_{0.1082}$
GLIMMIX (AR-1)	$0.0932^{\mathbf{3}}_{0.0063}$	$9.1258^{\mathbf{4}}_{0.3857}$	$0.2159^{\mathbf{5}}_{0.0136}$
NLMIXED	$0.0335^{\mathbf{6}}_{0.0060}$	$3.2856_{0.1699}$	$1.4055_{0.1175}$

Number of non-convergence outcomes from GLIMMIX procedures are tabulated as follows: **1.** 173; **2.** 7; **3.** 174; **4.** 1; **5.** 174; **6.** 7.

Distribution of the Random-Effect Estimators

Figure: Histogram of the random slope estimator from the binary data sets with $N = 100$ and $\rho = 0.7$. The true correlation structure of the data set is AR(1), and the estimators are obtained by the mixed-effect QIF method with AR(1) working correlation. The solid line provides the random-effects density function generated from the true Beta distribution.



A Binary Data Example

- An observational study, non-surgical periodontal treatment versus tooth loss
- 722 subjects, 7-year follow up
- Unbalanced data
- 550 (77%) patients have at least 5 years of follow up
- The response variable is tooth loss, a binary variable
- Apply a random-intercept model, the heterogeneity of patients is modeled as random effects
- To check whether the random-intercept model assumption is satisfied, the Chi-squared goodness-of-fit test (Qu et al., 2000) can be applied

Covariates

- Non-surgical periodontal treatment:
  - 1 if the patient continuously receives the treatment for all three years; and 0 otherwise
- Other covariates:
  - Gender
  - Age
  - Variables measuring the health condition of the teeth
    - Number of teeth (Teeth)
    - Number of diseased sites (Sites)
    - Mean pocket depth of diseased sites (Pddis)
    - Mean pocket depth of all sites (Pdall)
    - Number of non-periodontal treatments (Dent)
    - Number of non-periodontal preventive procedures (Prev)
    - Number of surgical treatments over the 3-year baseline period (Surg)
- The logistic model is

$$
\begin{aligned}
\texttt{logit}(\mu_{ij}^{b}) \quad = \quad & \beta_0 + b_i + \beta_1 \texttt{Gender}_{ij} + \beta_2 \texttt{Age}_{ij} + \beta_3 \texttt{Teeth}_{ij} + \beta_4 \texttt{Sites}_{ij} + \beta_5 \texttt{Pddis}_{ij} \\
& + \beta_6 \texttt{Pdall}_{ij} + \beta_7 \texttt{Surg}_{ij} + \beta_8 \texttt{Dent}_{ij} + \beta_9 \texttt{Prev}_{ij} + \beta_{10} \texttt{Nonsurg}_{ij}
\end{aligned}
$$

Comparison of PQL, QIF and GLIMMIX

Table: Comparison of the Mixed-Effect QIF and Other Approaches for the Periodontal Data

X	QIF_{ind}	QIF_{CS}	QIF_{AR}	PQL	GLMM_{ind}	GLMM_{AR}	NLM
Int	-7.1549	-7.4769	-8.1300	-8.0824	-9.3476	-9.5602	-6.8281
s.e.	1.3630	1.3476	1.3637	1.6265	1.7433	1.7804	1.5600
z-value	-5.2492	-5.5482	-5.9615	-4.9691	-5.3620	-5.3697	-4.3770
Gender	0.2257	0.2138	0.2409	0.2383	0.2317	0.2387	0.2526
s.e.	0.1522	0.1530	0.1589	0.1720	0.1766	0.1802	0.1588
z-value	1.4828	1.3974	1.5155	1.3865	1.3120	1.3246	1.5907
Age	0.0168	0.0175	0.0152	0.0202	0.0279	0.0291	0.0173
s.e.	0.0105	0.0104	0.0108	0.0123	0.0128	0.0131	0.0114
z-value	1.5948	1.6781	1.4072	1.6427	2.1772	2.2305	1.5109
Teeth	-0.0334	-0.0325	-0.0177	-0.0353	-0.0388	-0.0406	-0.0440
s.e.	0.0246	0.0241	0.0242	0.0271	0.2767	0.0282	0.0254
z-value	-1.3591	-1.3518	-0.7295	-1.3010	-1.4051	-1.4385	-1.7323
Sites	0.0024	0.0025	-0.0042	-0.0005	-0.0042	-0.0048	0.0032
s.e.	0.0097	0.0099	0.0090	0.0102	0.0105	0.0107	0.0098
z-value	0.2468	0.2555	-0.4684	-0.0524	-0.4029	-0.4440	0.3301

Cont'd

X	QIF_{ind}	QIF_{CS}	QIF_{AR}	PQL	GLMM_{ind}	GLMM_{AR}	NLM
Pddis	0.2689	0.3469	0.1948	0.2719	0.2944	0.2899	0.2587
s.e.	0.1864	0.1790	0.1904	0.2293	0.2370	0.2418	0.2124
z-value	1.4428	1.9377	1.0232	1.1866	1.2422	1.1989	1.2180
Pdall	0.4644	0.3960	0.7792	0.6425	0.8465	0.8885	0.4832
s.e.	0.3880	0.3946	0.3626	0.4200	0.4329	0.4423	0.4020
z-value	1.1968	1.0035	2.1489	1.5292	1.9554	2.0088	1.2020
Surg	-0.1377	0.0039	-0.1636	-0.0932	-0.1020	0.1304	-0.1087
s.e.	0.2741	0.2336	0.2863	0.2901	0.2019	0.2034	0.2790
z-value	-0.5024	0.0168	-0.5716	-0.3213	-0.5052	0.6411	-0.3896
Dent	0.1074	0.1132	0.1158	0.1205	0.1353	0.1365	0.1172
s.e.	0.0083	0.0082	0.0086	0.0080	0.0061	0.0061	0.0084
z-value	12.9164	13.8498	13.4963	15.0844	22.3636	22.4433	13.9126
Prev	0.0404	0.0271	0.0169	0.0353	0.0381	0.0395	0.0363
s.e.	0.1349	0.1353	0.1398	0.1500	0.0988	0.0990	0.1378
z-value	0.2992	0.2004	0.1207	0.2420	0.3856	0.3988	0.2636
Nonsurg	-0.2360	-0.2037	-0.2149	-0.2207	-0.1995	-0.2041	-0.2266
s.e.	0.1500	0.1504	0.1577	0.1767	0.1839	0.1876	0.1632
z-value	-1.5732	-1.3548	-1.3624	-1.2500	-1.0848	-1.0880	-1.3885

- In general, the standard errors of the conditional QIF are smaller than the PQL

Discussion

The advantages of the new approach:

- Incorporates both serial correlation from repeated measurements and heterogeneous variation from individuals
- Does not require the distribution assumption for random effects
- Does not require specifying the likelihood
- Does not need to estimate the unknown variance components or nuisance parameters associated with correlations

Discussion (Cont'd)

- Provides consistent and asymptotic normality for the fixed-effects estimator
- Outperforms the PQL, GLIMMIX, NLMIXED approaches when serial correlation is introduced, especially for binary response data
- Computationally fast even if the dimension of the random-effects parameters increases as the sample size increases
- GLIMMIX procedure tends to have a convergence problem

Thanks

- Annie Qu and Cindy Tsai
- Organizers of the workshop.
- BIRS staff and the cooks.
- Thank you for the attention.