Criteria for Working-Correlation-Structure Selection in GEE: Assessment via Simulation

Author(s): Lin-Yee Hin, Vincent J. Carey and You-Gan Wang

Source: *The American Statistician*, Nov., 2007, Vol. 61, No. 4 (Nov., 2007), pp. 360-364

Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association

Stable URL: https://www.jstor.org/stable/27643940

Criteria for Working–Correlation–Structure Selection in GEE: Assessment via Simulation

Lin-Yee HIN, Vincent J. CAREY, and You-Gan WANG

Efficiency of analysis using generalized estimation equations is enhanced when intracluster correlation structure is accurately modeled. We compare two existing criteria (a quasi-likelihood information criterion, and the Rotnitzky–Jewell criterion) to identify the true correlation structure via simulations with Gaussian or binomial response, covariates varying at cluster or observation level, and exchangeable or AR(1) intracluster correlation structure. Rotnitzky and Jewell's approach performs better when the true intracluster correlation structure is exchangeable, while the quasi-likelihood criteria performs better for an AR(1) structure.

KEY WORDS: Clustered data; Correlation structure; Generalized estimating equation.

1. INTRODUCTION

The efficiency of clustered data modeling via generalized estimating equations (GEE) proposed by Zeger and Liang (1986) is enhanced when the intracluster correlation structure is accurately modeled. Inaccurate specification of the intracluster correlation structure may affect the efficiency of parameter estimation without impairing consistency (Rotnitzky and Jewell 1990; Wang and Carey 2003).

For inferential purposes, nonparametric modeling of the intracluster correlation by specifying unstructured correlation as the working correlation structure in GEE modeling will likely suffice for comparative inference when the sample size is sufficiently large. However, if the GEE model developed is to be used as a population-averaged model for estimation or prediction, then an accurate parametric modeling of the intracluster correlation structure is desirable to reduce the potential error in estimation or prediction due to improved efficiency of parameter estimation.

Using the fact that the asymptotic distribution of a modified working Wald statistic and the working score test statistic are linear combinations of independent χ_1^2 random variables, Rotnitzky and Jewell (1990) described an approach to heuristically appraise the adequacy of the assumed intracluster correlation matrix in the process of modeling using GEE. Pan (2001) proposed a modification of the Akaike information criterion, called the "quasi-likelihood under the independence model criterion" (QIC(R)), to aid in choosing the best intracluster correlation structure for a given combination of covariates.

Since these two methods represent competing diagnostic tools that aid identification of the true intracluster correlation structure, it is important to compare how their performance varies under different data characteristics, including the nature of response variable, for example, Gaussian or binomial, and the true intracluster correlation structure, for example, exchangeable or AR(1), as well as the magnitude of such a correlation, as well as the nature of the covariates; for example, cluster-level, observation-level, or a mixture of both. The knowledge of their performance profiles can reveal their strengths and limitations, which will aid in their application and interpretation in clustered data analysis. In this article, we assess the performance profile of these two methods via simulation.

This article is organized as follows: Section 2 summarizes the main results in Zeger and Liang (1986), and describes the intracluster correlation structure selection technique proposed by Rotnitzky and Jewell (1990) and Pan (2001). Section 3 describes the simulation study performed to assess their performance and reports the results. Section 4 provides some concluding remarks.

2. NOTATION AND THEORETICAL BACKGROUND

Let there be n clusters in the dataset, and let t ($t = 1, \ldots, m_i$) index observation times within the ith cluster ($i = 1, \ldots, n$). Data from each cluster are represented as $\mathbf{y}_i$, an $m_i \times 1$ response vector, and $\mathbf{x}_i$, an $m_i \times p$ covariate matrix with tth row denoted $\mathbf{x}_{it}$. Here p will denote the number of covariates including, if one is present, an intercept term. Let β denote a p-vector of regression parameters, and let $g(\cdot)$ denote a link function in the family of generalized linear models. We define $g[E(y_{it})] = \eta_{it}$ where $\eta_{it} = \mathbf{x}_{it}\beta$ for the observation t ($t = 1, \ldots, m_i$) in cluster i, and we will use $\mu_i = g^{-1}(\mathbf{x}_i\beta)$ to denote the m_i-vector of mean response values for the ith cluster. Let $v(\cdot)$ denote a variance function (mean–variance relation) in the family of generalized linear models. Then we will use $\mathbf{A}_i$ to denote the $m_i \times m_i$ diagonal matrix of marginal variances for the ith cluster, with t element $v(\mu_{it})$.

The GEE approach (Liang and Zeger 1986) estimates $\boldsymbol{\beta}$ through solving the estimating equation

$$\sum_i \mathbf{D}_i^T \mathbf{V}_i^{-1}(\mathbf{y}_i - \boldsymbol{\mu}_i) = 0, \tag{1}$$

Lin-Yee Hin is a medical practitioner. Current address is 207, Yat Tung Shopping Center, Tung Chung, Hong Kong (E-mail: *lyhin@netvigator.com*). Vincent J. Carey is Associate Professor of Medicine (Biostatistics), Harvard Medical School, Channing Laboratory, Birmingham and Women's Hospital, 181 Longwood Avenue, Boston, MA 02115 (E-mail: *stvjc@channing.havard.edu*). Dr. Carey's work was supported in part by NIH Grant NHLBI HL 40619. You-Gan Wang is Senior Science Leader, CSIRO Mathematical and Information Sciences, Australia. Current address is CSIRO Mathematical and Information Sciences, Long Pocket Laboratories, 120 Meiers Road, Indooroopilly Queensland 4068, Australia (E-mail: *You-Gan.Wang@csiro.au*). The authors are grateful to the Editor, Associate Editor, and two referees for constructive and critical suggestions that led to a much improved article.

  DOI: 10.1198/000313007X245122

where $\mathbf{D}_i$ is the $m_i \times p$ matrix with (stacked) blocks $\partial \mu_i / \partial \beta$, and $\mathbf{V}_i = \mathbf{A}_i^{1/2} \mathbf{R}_i(\rho) \mathbf{A}_i^{1/2} / \phi$, with ϕ the dispersion parameter. The working correlation matrix for cluster i is denoted by $\mathbf{R}_i(\rho)$, an $m_i \times m_i$ square matrix depending on some q-vector of correlation parameters ρ.

Let $\mathbf{D}$ and $\mathbf{V}$ denote conforming blocked matrices constructed from the $\mathbf{D}_i$ and $\mathbf{V}_i$ defined above. Then according to Liang and Zeger (1986), the asymptotic covariance for the p covariates is given by a $p \times p$ matrix

$$V_\beta = \lim_{n\to\infty} n(\mathbf{D}^T \mathbf{V}^{-1} \mathbf{D})^{-1} \times \{\mathbf{D}^T \mathbf{V}^{-1} (\mathbf{Y} - \widehat{\boldsymbol{\mu}})(\mathbf{Y} - \widehat{\boldsymbol{\mu}})^T \mathbf{V}^{-1} \mathbf{D}\} \times (\mathbf{D}^T \mathbf{V}^{-1} \mathbf{D}_\beta)^{-1}, \quad (2)$$

where $\widehat{\boldsymbol{\mu}} = g^{-1}(\widehat{\boldsymbol{\eta}})$ with $\widehat{\boldsymbol{\eta}}$ being η evaluated at convergence.

2.1 Rotnitzky–Jewell's Criterion

According to Rotnitzky and Jewell (1990, sec. 4), we have

$$\mathbf{Q}_0 = n^{-1}(\mathbf{D}^T \mathbf{V}^{-1} \mathbf{D}) \quad (3a)$$

$$\mathbf{Q}_1 = n^{-1}\{\mathbf{D}^T \mathbf{V}^{-1} (\mathbf{Y} - \widehat{\mu})(\mathbf{Y} - \widehat{\mu})^T \mathbf{V}^{-1} \mathbf{D}\} \quad (3b)$$

with

$$\mathbf{Q} = \mathbf{Q}_0^{-1} \mathbf{Q}_1 \quad (4)$$

The adequacy of the working correlation matrix was examined by Rotnitzky and Jewell (hereafter RJ), defining $C_1 = \mathrm{tr}(\mathbf{Q})/p$ and $C_2 = \mathrm{tr}(\mathbf{Q}^2)/p$ where p is the rank of the model.

When the working correlation structure is correctly specified, $\mathbf{Q}$ should be close to an identity matrix, and hence both C_1 and C_2 should be close to 1 (Wang and Carey 2004). When comparing different candidate models, we may choose the model with (C_1, C_2) values closer to (1,1) using the Euclidean distance. Specifically, we can define this criteria based on the following measure

$$\mathrm{RJ}(R) = \surd\{(1 - C_1)^2 + (1 - C_2)^2\}. \quad (5)$$

The value of $\mathrm{RJ}(R)$ can be used as a criterion for selecting the most appropriate working correlation model among all the candidate models.

2.2 Quasi-Likelihood Under the Independence Model Criterion

Following the model specification of $E(y) = \mu$ and $\mathrm{var}(y) = \phi v(\mu)$, the log quasi-likelihood function is the path integral (McCullagh and Nelder 1989):

$$\mathcal{Q}(\mu; y) = \int_y^\mu \frac{y - t}{\phi v(t)} dt. \quad (6)$$

For independent observations, functional forms of $\mathcal{Q}$ for the members of the family of generalized linear models are available in Table 9.1 of McCullagh and Nelder (1989).

Let M_1 be the true model and M_2 the candidate model for a set of clustered data. Pan (2001) expressed the separation between M_1 and M_2 using the notion of Kullback–Leibler distance (Kullback and Leibler 1951), and subsequently obtains its approximation using Taylor's series expansion up to the second-order partial derivative. By ignoring the first-order partial derivative term that is difficult to estimate (Pan 2001), this approximation of the Kullback–Leibler distance, defined as $\mathrm{QIC}(R)$, can be expressed as

$$\mathrm{QIC}(R) = -2 \times \mathcal{Q}(g^{-1}(\widehat{\eta})) + 2 \times \mathrm{tr}(\mathbf{Q}_{0,I} \mathbf{V}_{\beta,R}), \quad (7)$$

Model-based is independent? I thought it was based on the working correlation structure

where $\mathbf{Q}_{0,I}$ is $\mathbf{Q}_0$ evaluated when the intra-cluster correlation structure is set at independence, and $\mathbf{V}_{\beta,R}$ is $\mathbf{V}_\beta$ evaluated with the intra-cluster correlation structure fixed at R, and the term involving $\mathcal{Q}$ is evaluated at convergence.

The presumed working intracluster correlation structure used in the GEE model fitting procedure yielding the lowest $\mathrm{QIC}(R)$, compared to those yield under different choices of working intracluster correlation structure, is the structure best describing the true underlying intracluster correlation structure (Pan 2001; Hardin and Hilbe 2003).

3. SIMULATION

We assess the performance of $\mathrm{QIC}(R)$ and $\mathrm{RJ}(R)$ in facilitating choice between exchangeable and AR(1) working correlation structures. Simulations contexts considered allow performance contrasts among situations involving predictors that vary only between-cluster, or only within-cluster, or vary both between and within clusters. In addition, we consider performance for both continuous and discrete responses.

We conduct a simulation study with 10,000 independent replications for each of the following settings. Each independent replication contains 100 balanced clusters ($n = 100$) of size 5 ($m_i = 5$), where x_1, x_2 are cluster level covariates, and x_{1t}, x_{2t} observation-level covariates, all generated at random from uniform distribution $U[0, 1]$. Cluster-level covariates take on the same value within a cluster, while observation-level covariate may take on different values within a cluster, and mixed-level covariates consist of one cluster-level and one observation-level covariate. Details of the simulation experiment are tabulated in Table 1. What is this?

Realization rates of 99.78% to 100% are achieved, for settings simulating Gaussian responses with $\rho = 0.8$ and mixed-level or observation-level covariates. However, the realization rates for settings simulating (1) cluster-level covariates, and (2) binary responses with $\rho = 0.8$ are as low as approximately 94%, hence 15,000 independent replications were performed for each of these settings, from which 10,000 independent replications are selected by random among the realized replications for analysis. Although the selections of 10,000 realized replications for the settings in question are performed at random, the possibility of selection bias cannot be excluded completely.

All computations are performed using R version 2.3.0 (R Development Core Team 2006), with GEE fitting performed using the `yags` library (Carey 2004). Gaussian random variable generation was performed using `MASS` library (Venables and Ripley 2002), and the binary random variables generated using `bindata` library (Leisch and Weingessel 2005). Model-based covariance and robust covariance are extracted from output of

Table 1. Frequencies of the intracluster correlation structure identified using RJ(R) from 10,000 independent replications, and their corresponding canonical links used in simulation.

Response type	Covariate level	ρ	True ICS	Selection by RJ(R) Ind	Ex	AR(1)	Selection by QIC(R) I:E	Ind	Ex	AR(1)	Canonical link
Gaus	Obs	0.4	Ex	0	8415	1585		946	6831	2223	$\mu_t = 0.3x_{1t} + 0.3x_{2t}$
Gaus	Obs	0.4	AR(1)	103	5278	4619		899	2079	7022	$\mu_t = 0.3x_{1t} + 0.3x_{2t}$
Gaus	Obs[2]	0.8	Ex	0	8670	1308		1207	6580	2191	$\mu_t = 0.3x_{1t} + 0.3x_{2t}$
Gaus	Obs[3]	0.8	AR(1)	0	5035	4963		873	2151	6974	$\mu_t = 0.3x_{1t} + 0.3x_{2t}$
Gaus	Mix	0.4	Ex	0	9317	683		1337	6495	2168	$\mu_t = 0.3x_{1t} + 0.3x_2$
Gaus	Mix	0.4	AR(1)	16	5696	4288		1249	2112	6639	$\mu_t = 0.3x_{1t} + 0.3x_2$
Gaus	Mix[1]	0.8	Ex	0	8803	1186		1606	6712	1671	$\mu_t = 0.3x_{1t} + 0.3x_2$
Gaus	Mix	0.8	AR(1)	0	5168	4830		1223	2241	6534	$\mu_t = 0.3x_{1t} + 0.3x_2$
Gaus	Cls	0.4	Ex	0	9762	238	7653			2347	$\mu = 0.5x_1 + 0.5x_2$
Gaus	Cls	0.4	AR(1)	14	6247	3739	3595			6405	$\mu = 0.5x_1 + 0.5x_2$
Gaus	Cls	0.8	Ex	0	8850	1150	8356			1644	$\mu = 0.5x_1 + 0.5x_2$
Gaus	Cls	0.8	AR(1)	0	5739	4261	3358			6642	$\mu = 0.5x_1 + 0.5x_2$
Bin	Obs	0.4	Ex	0	9610	390		922	6840	2238	$\text{logit}(\mu_t) = 0.62x_{1t} + 0.62x_{2t}$
Bin	Obs	0.4	AR(1)	34	5593	4373		877	2075	7048	$\text{logit}(\mu_t) = 0.62x_{1t} + 0.62x_{2t}$
Bin	Obs	0.8	Ex	0	7847	2153		1502	7721	777	$\text{logit}(\mu_t) = 0.1 + 0.1x_{1t}$
Bin	Obs	0.8	AR(1)	0	5348	4652		936	1806	7258	$\text{logit}(\mu_t) = 0.1 + 0.1x_{1t}$
Bin	Mix	0.4	Ex	0	9708	292		1206	6753	2041	$\text{logit}(\mu_t) = 0.6x_{1t} + 0.6x_2$
Bin	Mix	0.4	AR(1)	4	6081	3915		1100	2037	6863	$\text{logit}(\mu_t) = 0.6x_{1t} + 0.6x_2$
Bin	Mix	0.8	Ex	0	9143	857		1544	7352	1104	$\text{logit}(\mu_t) = 0.1 + 0.05x_{1t} + 0.05x_2$
Bin	Mix	0.8	AR(1)	0	5622	4378		966	1631	7403	$\text{logit}(\mu_t) = 0.1 + 0.05x_{1t} + 0.05x_2$
Bin	Cls	0.4	Ex	0	9843	157	8382			1618	$\text{logit}(\mu) = 1.2x_1 + 1.2x_2$
Bin	Cls	0.4	AR(1)	8	6498	3494	3299			6701	$\text{logit}(\mu) = 1.2x_1 + 1.2x_2$
Bin	Cls	0.8	Ex	0	10000	0	9796			204	$\text{logit}(\mu) = 0.1 + 0.2x_1$
Bin	Cls	0.8	AR(1)	0	7104	2896	1900			8100	$\text{logit}(\mu) = 0.1 + 0.2x_1$

NOTE: Gaus and Bin refer to response being Gaussian or binomial respectively. Cls and Obs refer to covariates being cluster-level or observation-level respectively. Mix refers to one covariate at cluster-level and one at observation-level. True ICS denotes true intracluster correlation structure (ICS) of simulated data. Ind, Ex and AR-1 refer ICS being independence, exchangeable and AR-1 respectively. I:E refers to ICS being Ind or Ex. ρ is the magnitude of intracluster correlation. [1]9,989 realized replications. [2]9,978 realized replications. [3]9,998 realized replications.

yags to calculate RJ(R), while the R script used to calculate QIC(R) is written based on the Stata macro described by Hardin and Hilbe (2003, p. 187). The results are tabulated in Table 1.

When the response is Gaussian and the true intracluster correlation structure is exchangeable, RJ(R) is successful in correct identification of the intracluster correlation structure in 84% to 97% of the replications. On the other hand, when the true intracluster correlation structure is AR(1), the correct identification rate is 37% to 49%.

In comparison, when the response is binary and the true intracluster correlation structure is exchangeble, the correct identification rates for RJ(R) range from 78% to 100%. When the true intracluster correlation structure is AR(1), the correct identification rate ranges between 28% to 46%.

When the response is Gaussian, the covariates are mixed-level or observation-level and the true intracluster correlation structure is exchangeable, QIC(R) demonstrates correct identification rates of 64% to 68% of the replications. When the response is binary, QIC(R) demonstrates a higher correct identification rate (67% to 77%) of true exchangeable correlation structure for the corresponding covariate levels.

QIC(R) demonstrates a correct identification rate of 64% to 70% in correctly detecting AR(1) correlation structure when the response is Gaussian. It demonstrates a higher correct detection rate for AR(1) structure of 67% to 74% when the response is binary.

For cluster-level covariates, QIC(R) demonstrates correct identification rate of 76% to 83% in detecting exchangeable structure, and a correct identification rate of 64% to 66% in detecting AR(1) structure. When the covariates are at cluster-level, the values of QIC(R) calculated under working independence and exchangeable correlation are identical among the realized replications. However, no coalescence in QIC(R) values are found between working independence and AR(1), or between exchangeable correlation and AR(1). This phenomenon does not occur when the covariates are either at mixed-level or observation-level.

That said, the usefulness of QIC(R) is not compromised by the QIC(R) coalescence between working independence and exchangeable structure. We can fit a preliminary model assuming exchangeable structure. If the estimated intracluster correlation value is virtually zero, then working independence will likely suffice for the modeling purpose. If the estimated intracluster correlation value is nonzero, then it will be necessary to find the suitable correlation structure from among AR(1) and other choices.

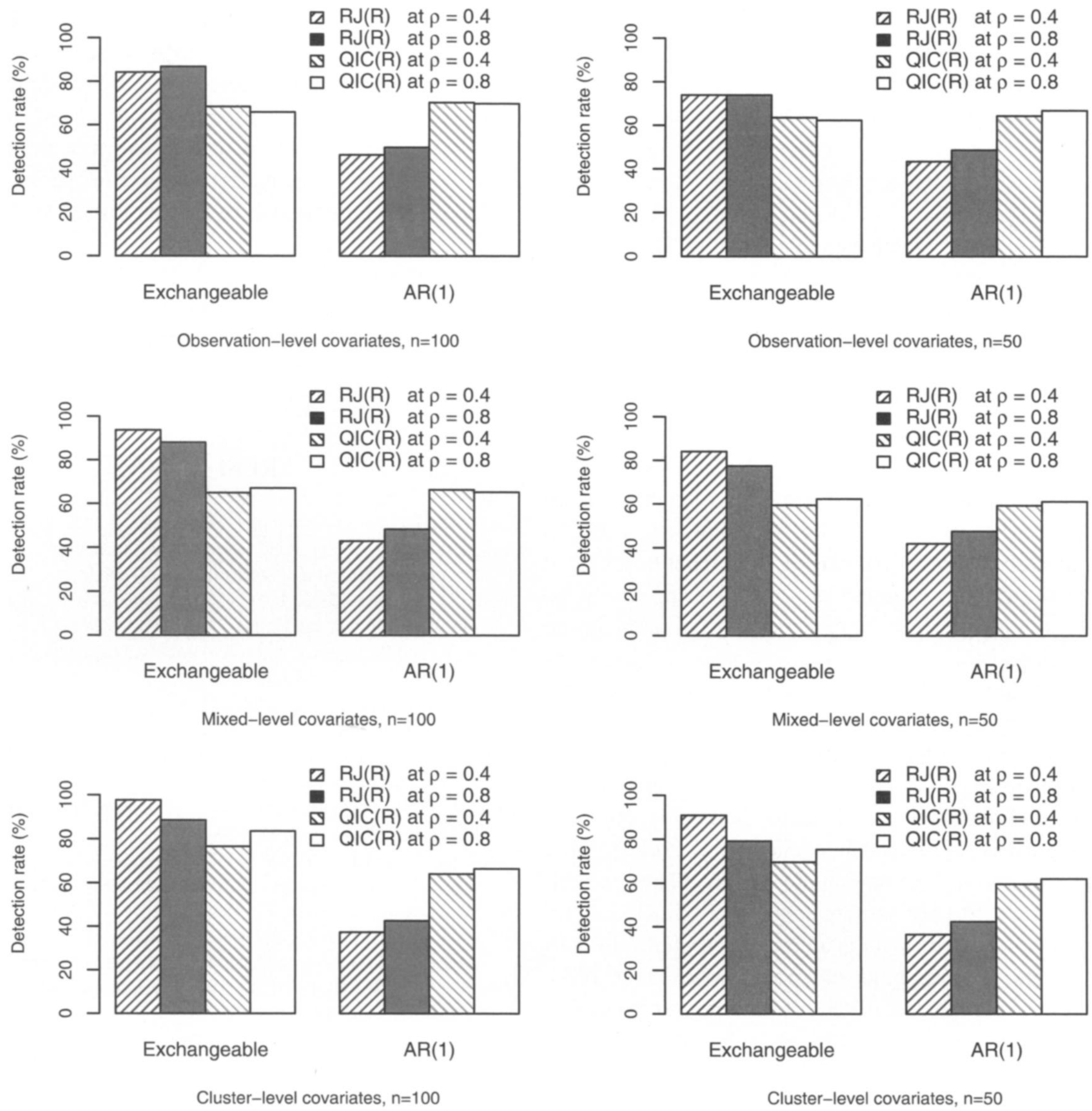

Figure 1. Intracluster correlation structure detection rate (%) of RJ(R) and QIC(R) when the simulated datasets contain either 100 balanced clusters ($n = 100$) or 50 balanced clusters ($n = 50$), each of size 5 ($m = 5$). 10,000 independent replications are generated for each setting. Ex refers to true intracluster correlation structure is exchangeable, while AR(1) is first-order auto regressive correlation.

For balanced design and cluster-level covariates, the magnitude of ρ in the exchangeable working model will not affect estimates of β regardless of whether link is identity. Thus, assuming the working independence model ($\rho = 0$) will produce the same $\hat{\beta}$ as that assuming exchangeable intracluster correlation. This can be mathematically explained as follows:

1. the mean (μ_i) and variance (ρ_i^2) do not change within the cluster;
2. $(1-\rho)R^{-1} = I - \rho\{1+(n-1)\rho\}^{-1}\mathbf{1}\mathbf{1}^{\mathrm{T}}$.

Suppose x_i is the $p \times 1$ vector, contribution from cluster i (with n observations) is

$$\mathbf{D}_i^{\mathrm{T}}\mathbf{V}_i^{-1}(\mathbf{y}_i - \mu_i) = \frac{1}{1+(n-1)\rho}\mathbf{D}_i^{T}\mathbf{A}_i^{-2}(\mathbf{y}_i - \mu_i). \qquad (8)$$

Here $\mathbf{D}_i^T = \partial\mu_i^T/\partial\beta$ is a $p \times n$ matrix. When n does not vary among the clusters, the constant $1+(n-1)\rho$ can be removed from the estimation equations in (8), and hence the GEE estimated under assumed exchangeable intracluster correlation is equivalent to the independence model. Also note that

$$\mathbf{D}_i^{\mathrm{T}}\mathbf{V}_i^{-1}\mathbf{D}_i = \frac{1}{1+(n-1)\rho}\mathbf{D}_i^{\mathrm{T}}\mathbf{A}_i^{-2}\mathbf{D}_i.$$

Therefore, it is not difficult to see that the corresponding sandwich estimator (2) will be free from the ρ values for the exchangeable working model. Hence, the value of QIC(R) calculated under working independence and that calculated under exchangeable correlation structure are the same.

In order to investigate the performance of RJ(R) and QIC(R) for a smaller sample size, we repeat the simulation for Gaussian response, the only difference being that each of the 10,000 independent replications contain 50 balanced clusters ($n = 50$) and each cluster of size 5 ($m = 5$). The results of the simulation (Gaussian response) for 50 balanced clusters ($n = 50$, $m = 5$)

are shown in Figure 1, and compared side by side with the results of simulation for 100 balanced clusters ($n = 100$, $m = 5$). It shows that the performance of RJ(R) and QIC(R) are similar for 50 balanced clusters and 100 balanced clusters.

4. DISCUSSION

From our simulation, it appears that neither QIC(R) nor RJ(R) can be regarded as a dominant criterion for working correlation structure selection. While RJ(R) is a tool of high sensitivity in terms of identifying exchangeable intracluster correlation, it has a low specificity. On the other hand, RJ(R) is a highly specific test for detecting AR(1) intracluster correlation, albeit with a low sensitivity. Therefore, if RJ(R) selects AR(1) (from among working independence, exchangeable, and AR(1) correlation structures) as the intracluster correlation structure, there is a good chance that the choice is correct. However, if RJ(R) selects exchangeable correlation instead, it may not be so.

As one of the referees pointed out, QIC(R) appears to be a more "robust" criteria since its detection rate remains in the 60% to 70% range for most scenarios regardless of the true intracluster correlation. On the other hand, RJ(R) performs better when the true intracluster correlation is exchangeable.

We have demonstrated the phenomenon of QIC(R) value coalescence in the case of balanced design and cluster-level covariates, and provided an explanation for it. In this situation, we can distinguish working independence from exchangeable correlation by examining the magnitude of intracluster correlation estimated via GEE, the former being approximately zero. Working independence, $\rho = 0$, is a conceptual entity. In real-life clustered data, working independence is nonexistent. If the true model is exchangeable, it is better to use the exchangeable correlation for inference because the model-based standard errors will be correct, and the independence model will provide much smaller (incorrect) standard errors. However, under this circumstances, the robust estimates for working independence and exchangeable correlation are both correct.

For mixed-level or observation-level covariates, QIC(R) demonstrates moderate sensitivity and specificity in identifying exchangeable intracluster correlation. In our simulation, for settings simulating binary responses with mixed or observation-level covariates, $\rho = 0.4$ or 0.8, and true intracluster correlation structure being exchangeable, the performance profile of QIC(R) is comparable to that described in Table 1 of Pan (2001).

Given the limitations of QIC(R) and RJ(R), more research is warranted to facilitate better modelling of the intracluster correlation structure for the GEE approach.

[Received March 2006. Revised August 2007.]

REFERENCES

Carey, V. (2004), *yags:Yet another GEE Solver*, R package version 4.0-2.

Kullback, S., and Leibler, R. A. (1951), "On Information and Sufficiency," *The Annals of Mathematical Statistics*, 22, 79–86.

Hardin, J. W., and Hilbe, J. M. (2003), *Generalized Estimating Equations*, Florida: Chapman and Hall.

Leisch, F., and Weingessel, A. (2005), *bindata: Generation of Artificial Binary Data*, R package version 0.9-12.

Liang, K.Y., and Zeger, S.L. (1986), "Longitudinal Data Analysis using Generalized Linear Models," *Biometrika*, 73, 13–22.

McCullagh, P., and Nelder, J.A. (1989), *Generalized Linear Models* (2nd ed.), London: Chapman and Hall.

Pan, W. (2001), "Akaike's Information Criterion in Generalized Estimating Equations," *Biometrics*, 57, 120–125.

Rotnitzky, A., and Jewell, N.P. (1990), "Hypothesis Testing of Regression Parameters in Semiparametric Generalized Linear Models for Cluster Correlated Data," *Biometrika*, 77, 485–497.

R Development Core Team (2006), *R. A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing.

Venables, W.N., and Ripley, B.D. (2002), *Modern Applied Statistics with S* (2nd ed.), New York: Springer.

Wang, Y.-G., and Carey, V. (2003), "Working Correlation Structure Misspecification, Estimation and Covariate Design: Implications for Generalised Estimating Equations Performance," *Biometrika*, 90, 29–41.

——— (2004), "Unbiased Estimating Equations from Working Correlation Models for Irregularly Timed Repeated Measures," *Journal of the American Statistical Association*, 99, 845–853.