Journal of the Korean Data & Information Science Society
2013, **24**(3), 651–658

http://dx.doi.org/10.7465/jkdi.2013.24.3.651
한국데이터정보과학회지

Quadratic inference functions in marginal models for longitudinal data with time-varying stochastic covariates[†]

Gyo-Young Cho[1] · Oyunchimeg Dashnyam[2]

[12]Department of Statistics, Kyungpook National University

Received 21 March 2013, revised 16 April 2013, accepted 21 April 2013

Abstract

For the marginal model and generalized estimating equations (GEE) method there is important full covariates conditional mean (FCCM) assumption which is pointed out by Pepe and Anderson (1994). With longitudinal data with time-varying stochastic covariates, this assumption may not necessarily hold. If this assumption is violated, the biased estimates of regression coefficients may result. But if a diagonal working correlation matrix is used, irrespective of whether the assumption is violated, the resulting estimates are (nearly) unbiased (Pan *et al.*, 2000).The quadratic inference functions (QIF) method proposed by Qu *et al.* (2000) is the method based on generalized method of moment (GMM) using GEE. The QIF yields a substantial improvement in efficiency for the estimator of β when the working correlation is misspecified, and equal efficiency to the GEE when the working correlation is correct (Qu *et al.*, 2000).In this paper, we interest in whether the QIF can improve the results of the GEE method in the case of FCCM is violated. We show that the QIF with exchangeable and AR(1) working correlation matrix cannot be consistent and asymptotically normal in this case. Also it may not be efficient than GEE with independence working correlation. Our simulation studies verify the result.

Keywords: FCCM assumption, GEE, longitudinal data, marginal model, QIF, time-varying stochastic covariates.

1. Introduction

In longitudinal analysis, there are several conditional expectations such as full covariate conditional mean $E(Y_{it} \mid X_{i1}, \ldots, X_{in_i})$, partly conditional mean $E(Y_{it}|\text{subset}\{X_{i1}, \ldots, X_{in_i}\}$ and cross-sectional mean $E(Y_{it} \mid X_{it})$. Then the first consideration for marginal regression analysis of longitudinal data with time-varying stochastic covariates is to determine the target of interest - that is whether full covariate conditional mean or certain partly conditional mean is of interest. The second issue is identification of valid and efficient estimation method.

[†] This research was supported by Kyungpook National University Research Fund, 2012.
[1] Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.
[2] Corresponding author: Ph. D student, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea. E-mail: d_ouynaa@yahoo.com

If one interest in cross-sectional association, we may choose the likelihood based methods such as linear and generalized linear mixed model and GEE method for marginal model. However there is implicit FCCM assumption that the conditional mean of the t^{th} response, given $X_{i1}, \ldots, X_{in_i}$, depends only on X_{it}.

$$E(Y_{it} \mid X_{i1}, \ldots, X_{in_i}) = E(Y_{it} \mid X_{it}) \tag{1.1}$$

With time-stationary covariates, this assumption necessarily holds since X_{it}=X_{ik} for all occasions $k \neq t$. Also, with time-varying covariates that are fixed by design of the study, the assumption also holds since values of the covariates at any occasion are determined a priori by study design and in a manner completely unrelated to the longitudinal response (Fitzmaurice *et al.*, 2004).

However, when a covariate is time-varying and stochastic, this assumption may not necessarily hold. At first, it is important to check the assumption made in (1.1), namely, that the conditional mean of the Y_{it}, given the entire time-varying covariate profile $X_{i1}, \ldots, X_{in_i}$, depends only on the covariate value at the t^{th} occasion, X_{it}.

For example, the assumption will be violated when the current value of Y_{it}, given X_{it} predicts the subsequent value of X_{it+1}. In this case $E(Y_{it} \mid X_{it}, X_{it+1}) \neq E(Y_{it} \mid X_{it})$ and X_{it+1} is said to confound the relationship between Y_{it} and X_{it}. For the AR(1) model for longitudinal data, the FCCM assumption is usually violated. Also, the model which was considered by Diggle *et al.* (2002) has AR(1) time-varying covariate and response variable of the depends on both current and lagged values of the covariate. For this model, $E\,(Y_{it} \mid X_{it-1}, X_{it}) \neq E\,(Y_{it} \mid X_{it})$ which means the FCCM assumption is violated.

In these cases, the regression parameters may be biased. Pepe and Anderson (1994) pointed out that when we use the GEE to do marginal regression, either a diagonal working correlation matrix should be used, or FCCM assumption needs to be validated. It means that they suggest using the independent working correlation matrix as a "safe" analysis choice. However the using working-independence correlation matrix in GEE guarantees consistency, but entails a serious loss of efficiency in many cases (Fitzmaurice, 1995).

The QIF method proposed by Qu *et al.* (2000) is an important and powerful alternative to the GEE and it does not require more assumptions than does the GEE method, but yields a substantial improvement in efficiency for the estimator of β when the working correlation is misspecified, and equal efficiency to the GEE when the working correlation is correct. The QIF is the method based on GMM introduced by Hansen (1982).

Lai and Small (2007) proposed another alternative GMM for the marginal regression analysis of longitudinal data with time-dependent covariates. They divided time-dependent covariates into three possible types and showed that their GMM has advantages over the GEE with independence working correlation for some type of time-dependent covariates.

In this paper, we interest in whether QIF has same problem as GEE or it can improve the results of the GEE method in the case of FCCM is violated. We show that the QIF with exchangeable and AR(1) working correlation matrix cannot be consistent and asymptotically normal in this case. Also it may not be efficient than GEE with independence working correlation. Our simulation studies verify the result.

This paper was organized as following: Section 2 shows that how the GEE estimator with independence working correlation can be consistent in the case of FCCM is violated and Section 3 introduces the review of QIF of Qu *et al.* (2000). Section 4 shows QIF for longitudinal data with time-varying stochastic covariates. Section 5 illustrates the results of simulation studies and final section is summary.

2. GEE estimation for longitudinal data with time-varying stochastic covariates

Consider the marginal model for the cross-sectional mean $\mu_{it} = E(Y_{it}|X_{it})$. The GEE assumes that the marginal mean μ_{it} is a function of the covariates through a link function g with $g(\mu_{it}) = X'_{it}\beta$, and the variance of Y_{it} is a function of the mean $\text{var}(Y_{it}) = \Phi V(\mu_{it})$, where Φ is the dispersion parameter.

The GEE estimator of the regression parameter μ_{it} is defined by

$$S_\beta(\beta,W)=\sum_{i=1}^{m}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T W_i\,(Y_i-\mu_i)=0 \tag{2.1}$$

where $W_i=V_i^{-1}=\left(A_i^{1/2}R_i(\alpha)A_i^{1/2}\right)^{-1}$ with A_i being the diagonal matrix of the marginal variances, $\text{var}(Y_{it})$ and $R_i(\alpha)$ being the working correlation matrix.

The GEE estimator is consistent if the estimating function is unbiased:

$$E(S_\beta(\beta,W))=0. \tag{2.2}$$

The estimating equations for the k-th covariates can be written as sums

$$S_\beta(\beta,W)=\sum_{i=1}^{N}\left[\sum_{t=1}^{n_i}\sum_{j=1}^{n_i}D_{itk}^T w_{itj}(Y_{it}-\mu_{it})\right] \tag{2.3}$$

where $\mu_{it} = E(Y_{it}|X_{it})$, $D_{itk}=\frac{\partial\mu_{it}}{\partial\beta_k}$ and w_{itj} is the t, j-th element of the weight matrix W_i. In order to ensure that $E(S_\beta(\beta,W))=0$ we can consider the expectation of each summand in (2.3),

$$\begin{aligned} E\left[D_{itk}^T w_{itj}\,(Y_{it}-\mu_{it})\right] &= E\left\{E\left[D_{itk}^T w_{itj}\,(Y_{it}-\mu_{it})\,\middle|\,X_{i1},\ldots,X_{in_i}\right]\right\}\\ &= E\left\{D_{itk}^T w_{itj}[E(Y_{it}|X_{i1},\ldots,X_{in_i})-\mu_{it}]\right\} \end{aligned} \tag{2.4}$$

If the FCCM condition is satisfied then $\mu_{it}= E(Y_{it}\mid X_{i1},\ldots,X_{in_i}) = E(Y_{it}\mid X_{it})$ and the estimating function is $E(S_\beta(\beta,W))=0$ and unbiased. On the other hand, if FCCM does not hold $\mu_{it}= E(Y_{it}\mid X_{it}) \neq E(Y_{it}\mid X_{i1},\ldots,X_{in_i})$, the estimating function will likely be biased and result in inconsistent estimates for the cross-sectional mean structure.

However, if diagonal weight matrix is used, then estimating equations for k-th covariates simplifies to

$$S_\beta(\beta,W)=\sum_{i=1}^{m}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T W_i\,(Y_i-\mu_i)=\sum_{i=1}^{N}\left[\sum_{t=1}^{n_i}D_{itk}^T w_{itt}(Y_{it}-\mu_{it})\right] \tag{2.5}$$

and the expectation of each summand

$$E[D_{itk}^T w_{itt}(Y_{it}-\mu_{it})]=E\left\{E\left[D_{itk}^T w_{itt}(Y_{it}-\mu_{it})\,\middle|\,X_{it}\right]\right\}=E\left\{D_{itk}^T w_{itt}[E(Y_{it}|X_{it})-\mu_{it}]\right\} \tag{2.6}$$

will have zero provided that μ_{it}=$E(Y_{it}|X_{it})$. It means that using independence working correlation in GEE can leads to consistent estimators.

However, for time-dependent covariates, Fitzmaurice (1995) shows that the independent working correlation can result in a substantial loss of efficiency for estimation of the coefficients associated with the time-dependent covariates and provides an example in which using the independent working correlation is only 60% efficient relative to the true correlation structure.

In the next section, we introduce QIF of Qu *et al.* (2000), new statistical methodology developed for the estimation and inference in longitudinal data analysis in marginal model.

3. QIF estimation for longitudinal data with time-varying stochastic covariates

3.1. Specification and estimation of QIF

The GEE solves the equation (2,1). The QIF is derived by observing that the inverse of the working correlation matrix can be approximated by a linear combination of several basis matrices:

$$R^{-1}=\sum_{i=1}^{m} a_i M_i \tag{3.1}$$

where $M_1 \cdots , M_m$ are known matrices and $a_1, \cdots , a_m$ are unknown constant.

Substituting (3.1) into (2,1), consider the following class of estimating functions:

$$S_\beta(\beta, W)=\sum_{i=1}^{N} \left(\frac{\partial \mu_i}{\partial \beta}\right)^T A_i^{-\frac{1}{2}} (a_1 M_1 + \ldots + a_m M_m) A_i^{-\frac{1}{2}} (Y_i - \mu_i). \tag{3.2}$$

Define the 'extended score' g_N to be

$$g_N(\beta) = \frac{1}{N}\sum_{i=1}^{N} g_i(\beta) = \frac{1}{N}\begin{pmatrix} \sum_{i=1}^{N} \left(\frac{\partial \mu_i}{\partial \beta}\right)^T A_i^{-\frac{1}{2}} M_1 A_i^{-\frac{1}{2}} (Y_i - \mu_i) \\ \cdot \\ \cdot \\ \cdot \\ \sum_{i=1}^{N} \left(\frac{\partial \mu_i}{\partial \beta}\right)^T A_i^{-\frac{1}{2}} M_m A_i^{-\frac{1}{2}} (Y_i - \mu_i) \end{pmatrix}. \tag{3.3}$$

The vector g_N contains more estimating equations than parameters, the GMM can be applied and define the quadratic inference function to be

$$Q_N(\beta) = g_N' C_N^{-1} g_N \tag{3.4}$$

where $C_N = (1/N^2) \sum_{i=1}^{N} g_i(\beta) g_i'(\beta)$. (3.5)

The quadratic inference function estimator $\widehat{\beta}$ is then defined to be

$$\widehat{\beta}=\text{argmin}_\beta \; Q_N(\beta) \; . \tag{3.6}$$

The QIF estimator is obtained with no need to estimate the nuisance correlation parameter. Hence, the QIF method does not rely on whether an appropriate estimation of the correlation parameter is available or not.

The QIF estimator has the usual large sample properties such as under suitable conditions, the GMM estimator is consistent, asymptotically normal. Also QIF estimator is equal or more efficient than the GEE estimator as shown Qu *et al.* (2000).

3.2. Comparison of QIF with GEE

As summarized Qu *et al.* (2000) and Song *et al.* (2009), when the working correlation structure is correctly specified, both the QIF and GEE are equally efficient. However, when the working correlation structure is misspecified, the QIF is more efficient than the GEE. Also the QIF is robust to outliers or contaminated data and provides both a goodness-of-fit test to validate the first moment marginal mean assumption and a model selection criterion to perform a stepwise regression analysis. These properties are either unavailable or difficult to establish in the GEE method.

However, the QIF has some limitations. The QIF depends on the availability of the basis matrices for a given correlation structure. Currently, the QIF is established only for four types of working correlation structures: Independence, exchangeability, AR(1) and unstructured. The current version of a QIF cannot handle the unequally spaced repeated measurements. Similar to the GEE, when missing data are present, the QIF works only under MCAR.

4. QIF for longitudinal data with time-varying stochastic covariates

As said in Section 2, the GEE estimation has limitation in that $\widehat{\beta}$ is not necessarily consistent when covariates vary over time (Pepe and Anderson, 1994, Davis, 2002). Specifically, (2.2) does not necessarily have expectation zero unless either we have valid FCCM assumption or we use working independence.

The QIF estimators can be more efficient than estimators from GEE even the working correlation matrix is misspecified in case FCCM is hold. Then in case of FCCM is violated, whether QIF has the same problem as GEE?

1. If we use diagonal working correlation in QIF, we can approximate as $R^{-1}= aI$. Then

$$g_N(\beta) = \frac{1}{N}\sum_{i=1}^{N} g_i(\beta) \approx \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T A_i^{-1}(Y_i-\mu_i) \tag{4.1}$$

In this case, dimension of the extended score is the same as the dimension of parameters, then minimizing QIF is equivalent to solving $g_N(\beta) = 0$ by method of moment.

This is GEE and in the case of independence working correlation, it can be unbiased for the longitudinal data with time-varying stochastic covariates. It means that if FCCM violated, using $R^{-1}= aI$ in QIF will result as GEE approach.

2. Exchangeable or AR(1) working correlation is used in QIF, it can be written as

$$R^{-1}=a_1I+a_2M_1.$$

(Qu *et al.,* 2000; Examples 1, 2).

In this case, GEE are

$$S_\beta(\beta,W)=\sum_{i=1}^{N}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T A_i^{-\frac{1}{2}}\left(a_1 I+a_2 M_1\right)A_i^{-\frac{1}{2}}\left(Y_i-\mu_i\right)=0 \tag{4.2}$$

and QIF will be

$$Q_N(\beta)=g_N^{'}C_N^{-1}g_N \tag{4.3}$$

where

$$g_N(\beta)=\frac{1}{N}\sum_{i=1}^{N}g_i(\beta)=\frac{1}{N}\begin{pmatrix}\sum_{i=1}^{N}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T A_i^{-1}\left(Y_i-\mu_i\right)\\ \sum_{i=1}^{N}\left(\frac{\partial\mu_i}{\partial\beta}\right)^T A_i^{-\frac{1}{2}}M_1A_i^{-\frac{1}{2}}\left(Y_i-\mu_i\right)\end{pmatrix} \tag{4.4}$$

$$C_N=\left(\frac{1}{N^2}\right)\sum_{i=1}^{N}g_i(\beta)g_i'(\beta) \tag{4.5}$$

As said previous, the QIF based on GMM method and the GMM estimator will be consistent in the case the moment conditions are correct.

To be moment conditions are correct,

$$g_N(\beta_o)=E\begin{pmatrix}\sum_{i=1}^{N}\left(\frac{\partial\mu_i(\beta_o)}{\partial\beta}\right)^T A_i^{-1}\left(Y_i-\mu_i(\beta_o)\right)\\ \sum_{i=1}^{N}\left(\frac{\partial\mu_i(\beta_o)}{\partial\beta}\right)^T A_i^{-\frac{1}{2}}M_1A_i^{-\frac{1}{2}}\left(Y_i-\mu_i(\beta_o)\right)\end{pmatrix}=0. \tag{4.6}$$

When FCCM is violated, we are assured $E(\sum_{i=1}^{N}(\frac{\partial\mu_{i(\beta_o)}}{\partial\beta})^T A_i^{-1}(Y_i-\mu_i(\beta_o)))=0$ as GEE in (2.6). But we cannot assured $E(\sum_{i=1}^{N}(\frac{\partial\mu_i(\beta_o)}{\partial\beta})^T A_i^{-1/2}M_1A_i^{-1/2}(Y_i-\mu_i(\beta_o)))=0$. It followed from (2.4) in case $\mu_{it}=E(Y_{it}\mid X_{it})\neq E(Y_{it}\mid X_{i1},\ldots,X_{in_i})$. That is, some of the moment conditions may be incorrect. So the QIF estimator cannot be consistent and asymptotically normal by Hansen's theorem (Hansen, 1982).

In summary, when we use the QIF in marginal regression for longitudinal data with time-varying stochastic covariates, either FCCM assumption needs to be validated or independence working correlation matrix should be used.

5. Simulation study

Setting 1. We consider the model of time-varying covariate vector X_i which is standardized AR-1 Gaussian process with autocorrelation parameter ρ. This setting was considered by Diggle et al. (2002)

$$Y_{it}=\gamma_o+\gamma_1X_{it}+\gamma_2X_{it-1}+b_i+e_{it} \tag{5.1}$$

$$X_{it}=\rho X_{it-1}+\epsilon_{it} \tag{5.2}$$

where $b_i, e_{it}, \epsilon_{it}$ are mutually independent, $b_i\sim N(0,1), e_{it}\sim N(0,1)$ and $\epsilon_{it}\sim N(0,1-\rho^2)$, $X_{i0}\sim N(0,1)$ and $X_{it}\sim N(0,1)$. This model implies

$$E(Y_{it}|X_{i1},\ldots X_{in})=\gamma_o+\gamma_1X_{it}+\gamma_2X_{it-1} \tag{5.3}$$

but yields the marginal mean

$$E(Y_{it}|X_{it}) = \beta_o + \beta_1 X_{it} \tag{5.4}$$

where $\beta_o = \gamma_0$ and $\beta_1 = \gamma_1 + \gamma_2 * \rho$. So FCCM assumption is violated.

We consider three estimators: 1) GEE/QIF with independence working correlation (GEE-ind/QIF-ind), 2) QIF using AR-1 working correlation (QIF-AR(1)), 3) QIF using exchangeable working correlation (QIF-exch).

Although current version of QIF can handle four kinds of working correlation structure as said in Section 3.2, QIF with unstructured working correlation use adaptive QIF which used variance matrix of responses instead of basic matrices. (Qu *et al.*, 2003). So unstructured working correlation is not compared.

We investigated several design features for this model and simulated 500 data sets each of which contained n=20, 50, 100 subject measures, 6, 10 time points and ρ=0.3, 0.5, 0.8 with $\gamma_0 = 0,\ \gamma_1 = 1,\ \gamma_2 = 1$.

We calculated bias of parameters and mean squared error (MSE) and simulated relative efficiency (SRE) as following:

$$MSE = E(\widehat{\beta}_o - \beta_0)^2 + E(\widehat{\beta}_1 - \beta_1)^2,\ SRE = \frac{\text{MSE of the GEE–ind estimator}}{\text{MSE of the QIF estimator}}\ .$$

We interested in the covariate distribution which have broad range of autocorrelation parameter values. Simulation results for ρ= 0.3, 0.5, 0.8 show that QIF results cannot be efficient than GEE-ind. The QIF-AR(1) working correlation is viewed better than the QIF-exch working correlation. This is because AR(1) working correlation close to the true correlation more than exchangeable working correlation. When autocorrelation parameter approached 1, the covariates are resembling to invariant covariate and some QIF results are becoming better. But we cannot say that the QIF-AR(1) result is better than the GEE-ind. The simulation result for ρ= 0.5 is shown as example.

Table 5.1 Bias, MSE, SRE for $\rho = 0.5$ for Setting 1

Working correlation		n=20		n=50		n=100	
		t=6	t=10	t=6	t=10	t=6	t=10
GEE-ind/QIF-ind	Biasβ_0	-0.005	-0.008	0.008	-0.003	-0.006	0.006
	Biasβ_1	-0.011	-0.001	-0.018	-0.015	-0.023	-0.015
	MSE	0.105	0.091	0.045	0.033	0.022	0.019
	SRE	1	1	1	1	1	1
QIF-AR(1)	Biasβ_0	-0.005	-0.008	0.008	-0.004	-0.006	0.006
	Biasβ_1	-0.009	-0.008	-0.018	-0.012	-0.023	-0.014
	MSE	0.112	0.107	0.047	0.036	0.022	0.019
	SRE	0.940	0.857	0.976	0.926	0.988	0.971
QIF-exch	Biasβ_0	-0.336	-0.008	0.034	0.017	-0.022	0.002
	Biasβ_1	0.357	0.008	-0.015	-0.014	-0.034	-0.014
	MSE	104.972	0.351	0.410	0.433	0.332	0.038
	SRE	0.001	0.260	0.110	0.077	0.065	0.493

Setting 2. In the next case, we consider the model (5.1) and (5.2), but covariate Xi which is AR(1) Gaussian process with variance $1/(1-\rho^2)$. Specifically, $X_{it} \sim N(0, 1/(1-\rho^2))$ and $\epsilon_{it} \sim N(0,1)$ for previous setting. For this case, the variance of the covariate X_i is more than 1. For this case we also could not found the best result from QIF than GEE-ind.

Table 5.2 Bias, MSE, SRE for $\rho = 0.5$ for Setting 2

Working correlation		n=20		n=50		n=100	
		t=6	t=10	t=6	t=10	t=6	t=10
GEE-ind/QIF-ind	Biasβ_0	0.003	0.008	0.002	-0.005	0.001	-0.006
	Biasβ_1	-0.032	-0.011	-0.016	-0.005	-0.024	-0.014
	MSE	0.101	0.071	0.039	0.03	0.019	0.016
	SRE	1	1	1	1	1	1
QIF-AR(1)	Biasβ_0	0.002	0.015	0.001	-0.005	0.001	-0.006
	Biasβ_1	-0.032	-0.02	-0.015	-0.002	-0.024	-0.014
	MSE	0.108	0.085	0.041	0.033	0.019	0.016
	SRE	0.931	0.838	0.967	0.928	0.997	0.976
QIF-exch	Biasβ_0	0.005	0.019	-0.013	-0.009	0.001	-0.011
	Biasβ_1	-0.031	-0.013	-0.018	-0.005	-0.025	-0.014
	MSE	0.278	0.341	0.286	0.151	0.115	0.054
	SRE	0.362	0.209	0.138	0.202	0.169	0.288

6. Summary

The QIF has several useful properties over the GEE in the case FCCM is satisfied. But for the marginal regression model for time-varying stochastic covariates, specifically when the FCCM assumption is violated, we say that the QIF do not provide consistent, asymptotically normal and efficient estimators over GEE with independence working correlation estimator.

We say that when we use the QIF in marginal regression for longitudinal data with time-varying stochastic covariates, either FCCM assumption needs to be validated or QIF with independence working correlation matrix should be used.

Our simulation result showed that QIF with AR(1), exchangeable working correlation cannot be efficient than GEE with independence working correlation as proved in Section 4.

References

Davis, C. S. (2002). *Statistical methods for the analysis of repeated measurements,* Springer-Verlag, New York.

Diggle, P. J., Heagerty, P., Liang, K-Y. and Zeger, S. L. (2002). *Analysis of longitudinal data*, Oxford University Press, New York.

Fitzmaurice, G. M, Liard, N.M. and Ware, J.H. (2004). *Applied longitudinal analysis,* Wiley, New York.

Fitzmaurice, G. M. (1995). A caveat concerning independence estimating equations with multivariate binary data. *Biometrics*, **51**, 309-317.

Lai, Tz. L. and Small, D. (2007). Marginal regression analysis of longitudinal data with time-dependent covariates: A generalised method of moments approach. *Journal of the Royal Statistical Society B*, **69**, 79-99.

Pan, W., Thomas, A. L. and John, E. C. (2000). Note on marginal linear regression with correlated response data. *The American Statistician.* **54**, 191-195.

Pepe, M. S. and Anderson, G. L. (1994). A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Communications in Statistics-Simulation*, **23,** 939-951.

Qu, A., Lindsay, B. G. and Li, B. (2000). Improving generalized estimating equations using quadratic inference functions. *Biometrika*, **87**, 823-836.

Qu, A. and Lindsay, B.G. (2003). Building adaptive estimating equation when inverse of covariance estimation is difficult. *Journal of the Royal Statistical Society B*, **65**, 127-142.

Song, P. X.-K., Jiang, Z., Park, E. J. and Qu, A. (2009) Quadratic inference functions in marginal models for longitudinal data. *Statistical Medicine*, **28**, 3683-3696