Science Journal of Applied Mathematics and Statistics
2014; 2(1): 20-25
Published online February 20, 2014 (http://www.sciencepublishinggroup.com/j/sjams)
doi: 10.11648/j.sjams.20140201.13

Robust covariance estimator for small-sample adjustment in the generalized estimating equations: A simulation study

Masahiko Gosho[1,*], Yasunori Sato[2], Hisao Takeuchi[3]

[1]Advanced Medical Research Center, Aichi Medical University, Aichi, Japan
[2]Clinical Research Center, Chiba University Hospital, Chiba, Japan
[3]Clinical Development Center, Asahikasei Pharma Corporation, Tokyo, Japan

Email address:
mgosho@aichi-med-u.ac.jp (M. Gosho), yasu@faculty.chiba-u.jp (Y. Sato), takeuchi.hp@om.asahi-kasei.co.jp(H. Takeuchi)

To cite this article:
Masahiko Gosho, Yasunori Sato, Hisao Takeuchi. Robust Covariance Estimator for Small-Sample Adjustment in the Generalized Estimating Equations: A Simulation Study.*Science Journal of Applied Mathematics and Statistics.*Vol. 2, No. 1, 2014, pp. 20-25.
doi: 10.11648/j.sjams.20140201.13

Abstract: The robust or sandwich estimator is common to estimate the covariance matrix of the estimated regression parameter for generalized estimating equation (GEE) method to analyze longitudinal data. However, the robust estimator would underestimate the variance under a small sample size. We propose an alternative covariance estimator to the robust estimator to improve the small-sample bias in the GEE method. Our proposed estimator is a modification of the bias-corrected covariance estimator proposed by Pan (2001, Biometrika88, 901—906) for the GEE method. In a simulation study, we compared the proposed covariance estimator to the robust estimator and Pan's estimator for continuous and binominallongitudinal responses for involving 10—50 subjects. The test size of Wald-type test statistics for the proposed estimator is relatively close to the nominal level when compared with those for the robust estimator and the Pan's approach.

Keywords: Bias, Binary Response, Continuous Response, Longitudinal Data, Test Size

1. Introduction

Correlated data are very commonly adopted in biomedical research in cases of repeated measurements on the same subject or clustered sampling. Reference [1] proposed the generalized estimating equations (GEE) method, which is one of the most popular ways to analyze these correlated data. The GEE method generally uses a robust covariance estimator (also known as sandwich or empirical covariance estimator) to estimate the covariance matrix of the regression coefficients. The covariance estimator is a consistent estimator, and is robust with respect to misspecification of the covariance matrix of the correlated data when the number of subjects is sufficiently large. However, the robust covariance estimator does not perform well for small samples. As it tends to be underestimated, the size of Wald-type tests for the regression coefficients could be substantially greater than the nominal level and the coverage probability of the corresponding confidence interval may be considerably below the nominal level for small samples even when the covariance matrix of the correlated data is correctly specified [2].

In practice, due to limited resources or limited size of the population in biomedical research, it is impossible to increase sample size. In pharmacological and toxicological researches on animals, the number of individuals per group would be about at most 10 and the total sample size is also about 50 individuals. In phase I studies, which are early conducted in the drug development process, the sample size is mostly 40—50.Of course, it is not easy to acquire the sufficient number of patients in clinical trials to develop orphan drugs. Even epidemiological studies, the sample size might be quite small. For example, an air pollution research reported by [3] was investigated on 16 children.

Two major approaches have been proposed to solve the small-sample issue. One is to use alternative test statistics or probability distribution in substitution for the Wald-type chi-squared test and the other is to directly correct the bias of the robust covariance estimator of the coefficient.

In the former case, [4, 5] used an *F*- or *t*-distribution rather than a chi-squared or normal distribution to test regression coefficients estimated by the GEE method. Further, [6] proposed the use of score statistics for the corresponding test and its modification in place of the Wald-type statistics. These researchers showed that in terms of the test size, their proposed approaches were superior to the Wald-type statistics based on the robust covariance estimator.

In the latter case, [2, 7, 8] proposed some alternative covariance estimators to improve the small-sample bias. Reference [9] compared the performance of the bias-corrected covariance estimators through a simulation study.

In this study, we propose a modification of the bias-corrected covariance estimator proposed by [8] for the GEE method to analyze the small-sample longitudinal data with continuous and binominalresponses, and we provide one of the options for researchers who routinely address research questions using data from small samples.

This paper is organized as follows. In Section 2, we provide the covariance estimator proposed by [8] and further, propose an alternative covariance estimator. In Section 3, we present the results of a simulation study to investigate the performance of the proposed and existing covariance estimators. Finally, Section 4 provides conclusions.

2. Bias-Corrected Robust Covariance Estimator

Let Y_{it} denote an outcome variable on subjects $i = 1, \ldots, K$ and observations $t = 1, \ldots, n_i$ for each subject. Also, assume that an$n_i \times p$ matrix of covariate values $\mathrm{X}_i = (x_{i1}, \ldots, x_{in_i})^T$ is adjoined to the outcome vector$\mathrm{Y}_i = (Y_{i1}, \ldots, Y_{in_i})^T$. To simplify notation, we suppose that $n_i = n$ as in [1].

The expected value and variance of the outcome variable are assumed to be $\mu_{it} = E(Y_{it}|x_{it}) = h^{-1}(x_{it}^T\beta)$ and $\mathrm{var}(Y_{it}|x_{it}) = \phi v(\mu_{it})$, respectively, where h is a specified link function, β is a regression parameter (p-vector) to be estimated, ϕ is a scale parameter, and v denotes a variance function that indicates the mean-variance relation. The working covariance matrix of Y_i, $\mathrm{V_i}$ is assumed to have the form $\phi \mathrm{A}_i^{1/2}\mathrm{R_i}(\alpha)\mathrm{A}_i^{1/2}$, in which $\mathrm{A}_i = \mathrm{diag}(v_{it})$ and $R_i(a)$ is the working correlation matrix parameterized by α, an association parameter (q-vector).

The GEE method identifies the estimator $\hat{\beta}$ of the regression parameter β as the solution to (1), substituting ϕ with a $K^{1/2}$-consistent estimator $\hat{\phi}(\mathrm{Y}, \beta)$ after replacing α with a $K^{1/2}$-consistent estimator $\hat{\alpha}(\mathrm{Y}, \beta, \phi)$.

$$\mathrm{U}(\beta) \equiv \sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{S}_i = 0, \tag{1}$$

whereD_i is an $n \times p$ matrix defined by $\mathrm{D}_i = \partial\mu_i/\partial\beta$, $\mathrm{V}_i = \phi\mathrm{A}_i^{1/2}\mathrm{R_i}(\alpha)\mathrm{A}_i^{1/2}$,$\mathrm{S_i} = \mathrm{Y}_i - \mu_i$,and $\mu_i = (\mu_{i1}, \ldots, \mu_{in})^T$.

The covariance matrix of$\hat{\beta}$ by the GEE method, which is referred to as the robust covariance, is given by

$$\mathrm{V_G} = \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{D}_i\right)^{-1} \times \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{cov}(\mathrm{Y}_i) \mathrm{V}_i^{-1} \mathrm{D}_i\right) \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{D}_i\right)^{-1} \tag{2}$$

According to [1], the covariance estimation $\mathrm{V_{RO}}$ of $\mathrm{V_G}$ in (2) can be obtained by replacing $\mathrm{cov}(\mathrm{Y}_i)$with $\mathrm{S}_i\mathrm{S}_i^T$. In addition, the model (or naive) covariance estimator of $\hat{\beta}$ is given by

$$\mathrm{V_{MO}} = \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{D}_i\right)^{-1}. \tag{3}$$

The robust covariance estimator $\mathrm{V_{RO}}$ using $\mathrm{S}_i\mathrm{S}_i^T$ is expected to underestimate the variance of $\hat{\beta}$ when the sample size K is small [2].$\mathrm{S}_i\mathrm{S}_i^T$is not an optimal estimator of $\mathrm{cov}(\mathrm{Y}_i)$, since it is neither consistent nor efficient, because it is based on the data from only one subject, i, as pointed out by [8].

Reference [8] proposed an alternative covariance estimator for $\hat{\beta}$ in (4) for small-sample adjustment:

$$\mathrm{V_{PA}} = \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{D}_i\right)^{-1} \times \left\{\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{A}_i^{\frac{1}{2}} \left(\frac{1}{K}\sum_{i=1}^{K} \mathrm{A}_i^{-\frac{1}{2}} \mathrm{S}_i\mathrm{S}_i^T \mathrm{A}_i^{-\frac{1}{2}}\right) \mathrm{A}_i^{\frac{1}{2}} \mathrm{V}_i^{-1} \mathrm{D}_i\right\} \times \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{D}_i\right)^{-1}. \tag{4}$$

As we can see in (4), [8] uses the following matrix as the estimator of $\mathrm{cov}(\mathrm{Y}_i)$:

$$\mathrm{W}_i = \mathrm{A}_i^{1/2} \left(\frac{1}{K}\sum_{i=1}^{K} \mathrm{A}_i^{-1/2} \mathrm{S}_i\mathrm{S}_i^T \mathrm{A}_i^{-1/2}\right) \mathrm{A}_i^{1/2}. \tag{5}$$

$\mathrm{W}_i$is a consistent estimator of $\mathrm{cov}(\mathrm{Y}_i)$; however, W_i is not an unbiased estimator for small K. Thus, we conventionally consider the use of

$$\mathrm{W}_i' = \mathrm{A}_i^{1/2} \left(\frac{1}{K-p}\sum_{i=1}^{K} \mathrm{A}_i^{-1/2} \mathrm{S}_i\mathrm{S}_i^T \mathrm{A}_i^{-1/2}\right) \mathrm{A}_i^{1/2} \tag{6}$$

as the estimator of $\mathrm{cov}(Y_i)$. Hence, a modified covariance estimator is given by

$$\mathrm{V_{PM}} = \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{D}_i\right)^{-1} \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{W}_i' \mathrm{V}_i^{-1} \mathrm{D}_i^T\right) \times \left(\sum_{i=1}^{K} \mathrm{D}_i^T \mathrm{V}_i^{-1} \mathrm{D}_i\right)^{-1}. \tag{7}$$

In practice, the covariance estimator $\mathrm{V_{PM}}$ can be obtained by replacing β, ϕ, α with their respective estimates in (7). Of course, $\mathrm{W}_i = \mathrm{W}_i'$ and $\mathrm{V_{PA}} = \mathrm{V_{PM}}$ when $K \gg p$and $K = \infty$. Note that $\mathrm{W}_i'$is not dependent on ϕ, just like W_i.

The modification is to use a degrees-of-freedom correction similar to the one conventionally used to obtain unbiased estimates of $\mathrm{cov}(\mathrm{Y}_i)$. $\mathrm{V_{PM}}$is more efficient than $\mathrm{V_{RO}}$and this result is simply proven by the discussion in[8].

Our proposed covariance estimator is slightly insight because $1/K$ in (4) proposed by [8] is only replaced with $1/(K-p)$ and it is well known as a bias adjustment in general linear models. However, the magnitude of the covariance estimates is meaningfully different between Pan's and our methods when K is small. Reference [2] have represented a simple-to-implement modification to the GEE robust estimator in (2), which consists of multiplying the GEE robust estimator by $K/(K-p)$, and evaluated the performance of the modified estimator. This idea may be based on [10] which have considered a similar modification to the heteroskedasticity-consistent estimator shown in [11].

Henceforth, the robust and model covariance estimators in (2) and (3), the bias-corrected covariance estimator in(4), and our proposed estimator in (7) are referred to as RO, MO, PA, and PM, respectively.

3. Simulation Study

3.1. Scenarios and Methods

The small-sample property of the proposed covariance estimator was evaluated by a simulation study. We assumed a randomized clinical trial with parallel group design, which each subject is assigned to either treatment group or control group. In this simulation study, the test size based on Wald z-statistics of β was investigated. The test size was defined as the proportion of times $|\hat{\beta}_j/\mathrm{SE}(\hat{\beta}_j)| \geq z_{0.975}(j = 0,\cdots,p-1)$ under a null hypothesis $H_0\colon \beta_j = \beta_{j(\mathrm{true})}$, with $z_{0.975}$ as the 97.5 th percentile of standard normal distribution, $\hat{\beta}$ as the GEE estimate, and SE denoting the standard error derived from RO, MO, PA, and PM. We focusedlargely on the comparison between the test sizes of PM and PA because the statistical test for the regression parameters is of clinical and statistical interests in randomized clinical trials, andour proposed estimator PM is a modification of the bias-corrected covariance estimator PA proposed by [8].

The simulation data were generated for six scenarios presented in Table 1. In Scenarios 1 to 3 (normal response), we assumed a multivariate normal distribution with mean μ_{it}, variance σ_{it}^2, and correlation coefficient $\rho_{itt'}$.The variance was$\sigma_{it}^2 = 1$ and the true correlation structures were exchangeable with $\rho_{itt'} = \rho = 0.2$ and 0.5 .In Scenarios 4 to 6 (binominal response), the binominal response was assumed to be a multivariate binominal distribution with mean μ_{it}[12].The true correlation structure was an exchangeable structure with the true correlation coefficient $\rho = 0.2$.

In the six scenarios, x_i was a treatment group (subject-between covariate) indicator; half the subjects had $x_i = 0$ and the other half had $x_i = 1$. z_{it} was a subject-within covariate indicator and an independent Bernoulli distribution, that is, $z_{it} = 0$ or 1 with probability $1/2$.

We correctly fitted the marginal mean model with $\beta_0 + \beta_1 x_i$ for Scenario 1, $\beta_0 + \beta_1 x_i + \beta_2 z_{it}$for Scenario 2, and $\beta_0 + \beta_1 x_i + \beta_2(t-1) + \beta_3 x_i(t-1)$ for Scenario 3 to the simulated normal response. In addition, we correctly fitted the marginal logistic model with $\exp(\beta_0 + \beta_1 x_i)$ for Scenario 4, $\exp(\beta_0 + \beta_1 x_i + \beta_2 z_{it})$ for Scenario 5, and $\exp(\beta_0 + \beta_1 x_i + \beta_2(t-1) + \beta_3 x_i(t-1))$ for Scenario 6 to the simulated binominal response. $\beta = (\beta_0, \beta_1)^T$ for Scenarios 1 and 4, $\beta = (\beta_0, \beta_1, \beta_2)^T$ for Scenarios 2 and 5, and $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^T$ for Scenarios 3 and 6 were estimated by the GEE method using the exchangeable correlation structure or AR(1) structure as the working correlation structure.

Further, we set the sample size K as 10, 16, 20, 30, 40, and 50 and the number of observations n as 5, 10, and 20 for each subject. Data generation was repeated 100,000 times in each scenario. The simulation study was performed by using SAS 9.3 (SAS institute Inc., Cary, NC).

Table 1.*Scenarios for generating simulation data*

Scenario	Data type	True mean structure	True regression parameter$\boldsymbol{\beta}_{(\mathrm{true})}$
1	Normal	$\mu_{it} = \beta_{0(\mathrm{true})} + \beta_{1(\mathrm{true})}x_i$	$(0,0)$
2	Normal	$\mu_{it} = \beta_{0(\mathrm{true})} + \beta_{1(\mathrm{true})}x_i + \beta_{2(\mathrm{true})}z_{it}$	$(0,0,0)$
3	Normal	$\mu_{it} = \beta_{0(\mathrm{true})} + \beta_{1(\mathrm{true})}x_i + \beta_{2(\mathrm{true})}(t-1) + \beta_{3(\mathrm{true})}x_i(t-1)$	$(0,0,0,0)$
4	Binominal	$\mathrm{logit}(\mu_{it}) = \beta_{0(\mathrm{true})} + \beta_{1(\mathrm{true})}x_i$	$(0,0)$
5	Binominal	$\mathrm{logit}(\mu_{it}) = \beta_{0(\mathrm{true})} + \beta_{1(\mathrm{true})}x_i + \beta_{2(\mathrm{true})}z_{it}$	$(0,0,0)$
6	Binominal	$\mathrm{logit}(\mu_{it}) = \beta_{0(\mathrm{true})} + \beta_{1(\mathrm{true})}x_i + \beta_{2(\mathrm{true})}(t-1) + \beta_{3(\mathrm{true})}x_i(t-1)$	$((0,0,0,0)$

3.2. Results

Figs. 1 to 6 present the results of the simulation study in terms of the test sizes based on Wald z-statistics of β_j in the cases where MO, RO, PA, and PM were used. The results are not reported for the intercept term β_0 since it is usually considered to be a nuisance parameter in the six scenarios.

Fig. 1 shows the test size of β_1 for Scenario 1when ρ =0.5, in which an exchangeable structure is correctly specified and AR(1) is misspecified as the working correlation structure. The test size for PM was rather closer to 0.05 than those for MO, RO, and PA at all times in this scenario. The test size for PM also increased as the sample size Kdecreased, much as with the other estimators. In fact, the test size for PM was somewhat inflated and fell within the range 0.082−0.088when $K = 10$. However, the size for PM was approximately 0.05 when $K \geq 30$. The size of the effect for the subject-between covariate may have been inflated even when PA was applied. The results are not reported for the test size when $\rho = 0.2$ since they are quite similar to that when $\rho = 0.5$.It is the same also at the time of Scenarios 2 and 3.

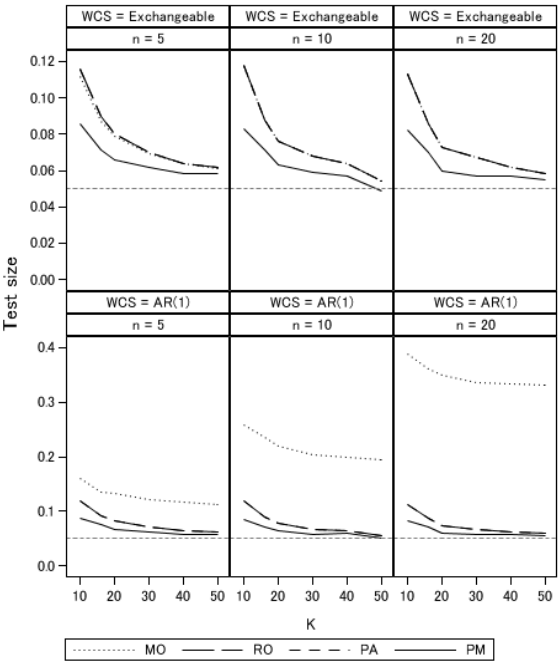

Figure 1.*Test size of Wald-type test statistics for* β_1 *rejecting hypothesis* H_0: $\beta_1 = 0$ *at a nominal 0.05 level in Scenario 1 (normal response). WCS, working correlation structure; MO, model covariance estimator; RO, robust covariance estimator; PA, covariance estimator by [8]; PM, proposed covariance estimator.*

Fig. 2 reports the test sizes of β_1 and β_2 when $n = 10$ for Scenario 2, in which an exchangeable structure is correctly specified and AR(1) is misspecified as the working correlation structure. The results are not reported for the test size when $n = 5, 20$and $\rho = 0.2$since they are quite similar to that when $n = 10$and $\rho = 0.5$. It is the same also at the time of Scenario 3. The test size for PM of β_1 corresponding to the treatment effect fell within the range $0.047 - 0.074$ when $n = 5, 10, 20$ and $\rho = 0.2, 0.5$. In addition, the test size for the PM of β_2 corresponding to the effect of the subject-within covariate was between 0.024 and 0.047 and PM was conservative when $K \leq 20$. On the other hand, the test size for the PA of β_2 was between 0.051 and 0.080when $K \leq 20$. The test sizes for all estimators except MO were slightly dependent on the specification of the working correlation structure.

Fig. 3 presents the test sizes of β_1, β_2, and β_3 when $n = 10$ and $\rho = 0.5$ for Scenario 3, in which an exchangeable structure is correctly specified and AR(1) is misspecified as the working correlation structure. The test size for PM of β_1 fell within the range $0.045{-}0.056$and was close to the nominal level 0.05when $n = 5,10,20$and $\rho = 0.2, 0.5$. In addition, the test sizes for the PM of β_2 and β_3 corresponding to the time variable and the interaction term between x_i and $t - 1$, respectively, were generally controlled by the nominal level.

Fig. 4 shows the test size of β_1 for Scenario 4, in which an exchangeable structure is correctly specified and AR(1) is misspecified as the working correlation structure. The test size for PM was rather closer to 0.05 than those for MO, RO, and PA at all times in this scenario. The test size for PM also increased as the sample size K decreased, much as with the other estimators. In fact, the test size for PM was somewhat inflated and fell within the range $0.064 - 0.076$ when $K = 10$. However, the size for PM was approximately 0.05 when $K \geq 40$. Further, the size for PM increased as the number of observations n increased when K was small. This property was similar to the other estimators. The test size for PA was quite similar to that for RO. The size of the effect for the subject-between covariate may have been inflated even when PA was applied.

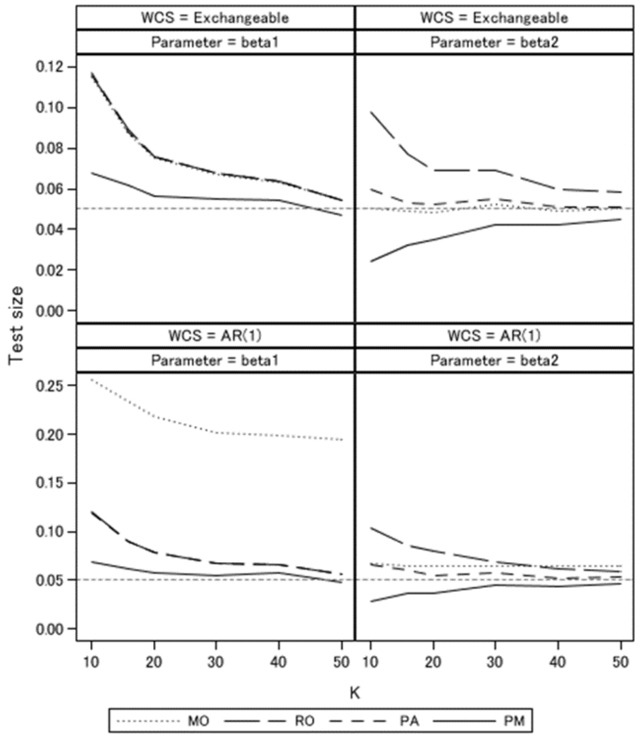

Figure 2.*Test size of Wald-type test statistics for* β_j *rejecting hypothesis* H_0: $\beta_j = 0$ *(j = 1, 2) at a nominal 0.05 level in Scenario 2 (normal response) when* $n = 10$ *and* $\rho = 0.5$. *WCS, working correlation structure; MO, model covariance estimator; RO, robust covariance estimator; PA, covariance estimator by [8]; PM, proposed covariance estimator.*

Fig. 5 reports the test sizes of β_1 and β_2 when $n = 10$ for Scenario 5, in which an exchangeable structure is correctly specified and AR(1) is misspecified as the working correlation structure. The results are not reported for the test size when $n = 5, 20$ since they are quite similar to that when $n = 10$. It is the same also at the time of Scenario 6. According to Fig. 5, the test size for PM of β_1 corresponding to the treatment effect fell within the range $0.049{-}0.063$when $n = 5, 10, 20$. In addition, the test size for the PM of β_2 corresponding to the effect of the subject-within covariate was between 0.021 and 0.047 and PM was conservative when$K \leq 20$. The test size for the PA of β_2 was between 0.049 and 0.063 and was somewhat controlled at the nominal size. This result was most like the discussion by [8]. The test sizes for all estimators except MO were slightly dependent on the specification of the working correlation structure.

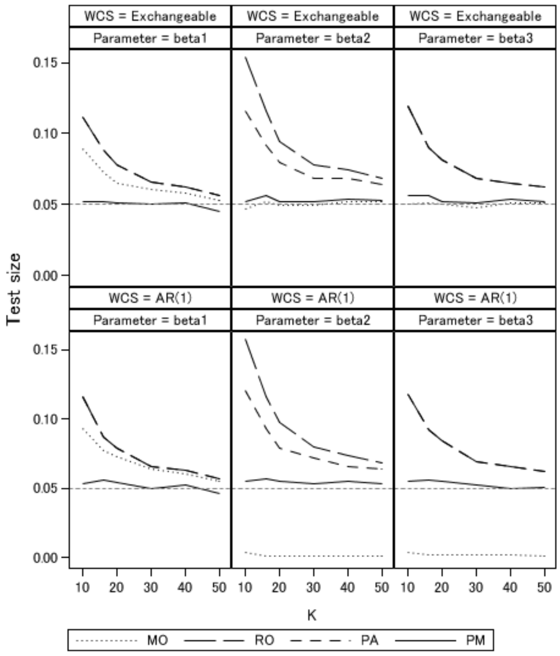

Figure 3.*Test size of Wald-type test statistics for β_j rejecting hypothesis H_0: $\beta_j = 0$ $(j = 1, 2, 3)$ at a nominal 0.05 level in Scenario 3 (normal response) when n =10 and ρ= 0.5. WCS, working correlation structure; MO, model covariance estimator; RO, robust covariance estimator; PA, covariance estimator by [8]; PM, proposed covariance estimator.*

Fig. 6 presents the test sizes of β_1, β_2, and β_3 when $n = 10$ for Scenario 6, in which an exchangeable structure is correctly specified and AR(1) is misspecified as the working correlation structure. The test size for PM of β_1 fell within the range 0.029−0.051 when $n = 5,10,20$. The size was smaller than the nominal level 0.05 when $K \leq 30$. In addition, the test sizes for the PM of β_2 and β_3 corresponding to the time variable and the interaction term between x_i and $t - 1$, respectively, were generally controlled by the nominal level 0.05.

Through the six scenarios, the sizes for PM of the treatment effect were relatively closer to the nominal size, especially for small K, compared with that for PA. We also reconfirmed the small-sample problem for RO even when the correlation structure was correctly specified.Further, our proposed estimator PM can be interpreted as a helpful and robust modification of PAto address small-sample issues because the test size for the proposed method wasnot particularly dependent on the number of observations for each subject, the correlation structure, and the distribution of the response variable.

4. Conclusions

We modified the covariance estimator by [8] to adjust the small-sample bias of the covariance estimator of the regression coefficients. The GEE robust covariance estimator tends to underestimate the variance under a small sample size; that is, the p value for the robust covariance estimator is small and the statistical inference is liberal.

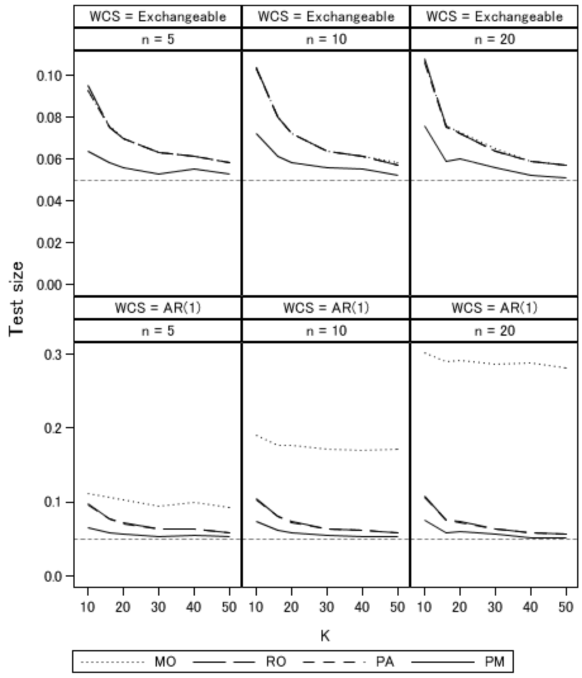

Figure 4.*Test size of Wald-type test statistics for β_1 rejecting hypothesis H_0: $\beta_1 = 0$ at a nominal 0.05 level in Scenario 4 (binominal response). WCS, working correlation structure; MO, model covariance estimator; RO, robust covariance estimator; PA, covariance estimator by [8]; PM, proposed covariance estimator.*

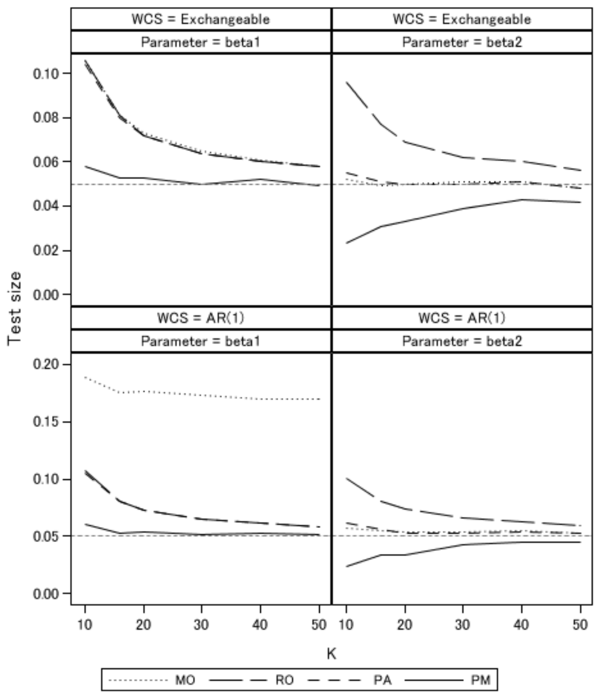

Figure 5.*Test size of Wald-type test statistics for $\boldsymbol{\beta_j}$ rejecting hypothesis H_0: $\boldsymbol{\beta_j} = \mathbf{0}$ (j = 1, 2) at a nominal 0.05 level in Scenario 5 (binominal response) when $\boldsymbol{n}$ =10. WCS, working correlation structure; MO, model covariance estimator; RO, robust covariance estimator; PA, covariance estimator by [8]; PM, proposed covariance estimator.*

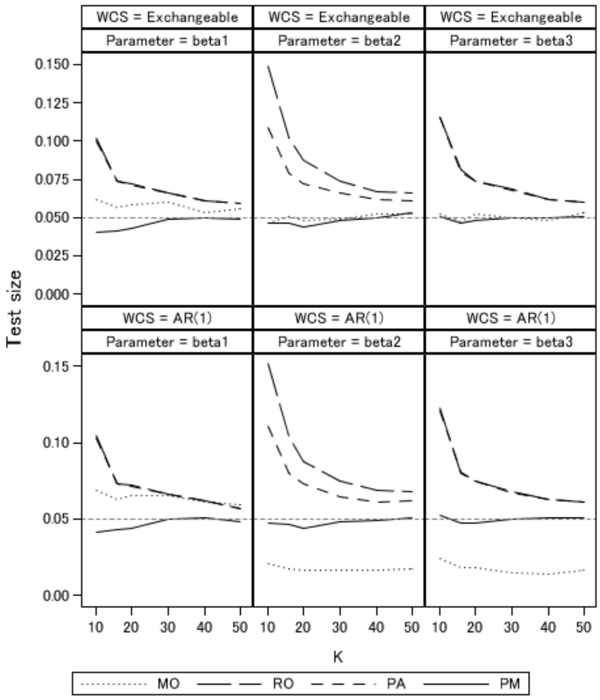

Figure 6.*Test size of Wald-type test statistics for* β_j *rejecting hypothesis* H_0: $\beta_j = 0$ $(j = 1, 2, 3)$ *at a nominal 0.05 level in Scenario 6 (binominal response) when* $n = 10$*. WCS, working correlation structure; MO, model covariance estimator; RO, robust covariance estimator; PA, covariance estimator by [8]; PM, proposed covariance estimator.*

The proposed covariance estimator avoids the bias of the robust variance estimator under the small sample size in many situations. The proposed method cannot completely control the test size below a nominal level; however, the method has relatively good performance compared with the existing methods.

Our estimator has a simple structure, like that of [8]. In the data analysis for small samples, we often confront the convergence problem and the impossibility of estimates for the regression parameters.The simplicity of the covariance estimator will be an important property for the small sample issues.

Acknowledgments

The authors wish to thank the reviewers for their helpful comments and suggestions, which led to substantial improvements in the article.

References

[1] Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linearmodels. *Biometrika*73:13—22.

[2] Mancl, L. A. and DeRouen, T. A. (2001). A covariance estimator for GEE with improvedsmall-sample properties. *Biometrics* 57:126—134.

[3] Hardin, J. and Hilbe, J. (2013). *Generalized Estimating Equations, 2nd edn.* Chapman and Hall, London.

[4] Fay, M. P. and Graubard, B. I. (2001). Small-sample adjustments for Wald-type testsusing sandwich estimators. *Biometrics* 57:1198—1206.

[5] Pan, W. and Wall, M. M. (2002). Small-sample adjustments in using the sandwich variance estimator in generalized estimating equations. *Statistics in Medicine*21:1429—1441.

[6] Guo, X., Pan, W., Connett, J., Hannan, P., and French, S.A. (2005). Small-sample performance of the robust score test and its modifications in generalized estimating equations. *Statistics in Medicine* 24:3479—3495.

[7] Kauermann, G. and Carroll, R. (2001). A note on the efficiency of sandwich covariancematrix estimation. *Journal of the American Statistical Association* 96:1387—1396.

[8] Pan, W. (2001). On the robust variance estimator in generalisedestimatingequations. *Biometrika*88:901—906.

[9] Lu, B., Preisser, J. S., Qaqish, B. F., Suchindran, C., Bangdiwala, S. I., andWolfson, M. (2007). A comparison of two bias-corrected covariance estimators for generalizedestimating equations. *Biometrics* 63:935—941.

[10] MacKinnon, J. and White, H. (1985). Some heteroskedasticity-consistent covariancematrix estimators with improved finite sample properties. *Journal of Econometrics* 29:305—325.

[11] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimatorand a direct test for heteroskedasticity. *Econometrika* 48:817—838.

[12] Qaqish, B. F. (2003). A family of multivariate binary distributions for simulatingcorrelated binary variables with specified marginal means and correlations.*Biometrika* 90:455—463.