

RESEARCH

Open Access



# A bias-reduced generalized estimating equation approach for proportional odds models with small-sample longitudinal ordinal data

Yukio Tada<sup>1\*</sup> and Tosiya Sato<sup>2</sup>

## Abstract

**Background** Longitudinal ordinal data are commonly analyzed using a marginal proportional odds model for relating ordinal outcomes to covariates in the biomedical and health sciences. The generalized estimating equation (GEE) consistently estimates the regression parameters of marginal models even if the working covariance structure is misspecified. For small-sample longitudinal binary data, recent studies have shown that the bias of regression parameters may result from the GEE and have addressed the issue by applying Firth's adjustment for the likelihood score equation to the GEE as if generalized estimating functions were likelihood score functions. In this manuscript, for the proportional odds model for longitudinal ordinal data, the small-sample properties of the GEE were investigated, and a bias-reduced GEE (BR-GEE) was derived.

**Methods** By applying the adjusted function originally derived for the likelihood score function of the proportional odds model to the GEE, we produced the BR-GEE. We investigated the small-sample properties of both GEE and BR-GEE through simulation and applied them to a clinical study dataset.

**Results** In simulation studies, the BR-GEE had a bias closer to zero, smaller root mean square error than the GEE with coverage probability of confidence interval near or above the nominal level. The simulation also showed that BR-GEE maintained a type I error rate near or below the nominal level.

**Conclusions** For the analysis of longitudinal ordinal data involving a small number of subjects, the BR-GEE is advantageous for obtaining estimates of the regression parameters of marginal proportional odds models.

**Keywords** Bias reduction, Marginal model, Categorical data, Penalization, Firth's adjustment

## Introduction

Longitudinal ordinal data are frequently collected in biomedical and health science studies. In such a study, each subject is observed repeatedly over a period, and ordinal responses of interest at each observation are recorded. Because repeated responses from the same subject are usually correlated, a straightforward application of generalized linear models for a single response variable to longitudinal data is not appropriate. There are numerous approaches for extending generalized linear models to

\*Correspondence:

Yukio Tada  
yukio\_tada@shionogi.co.jp

<sup>1</sup> Biostatistics Center, Shionogi & Co., Ltd., Osaka, Japan

<sup>2</sup> Department of Biostatistics, Kyoto University School of Public Health, Kyoto, Japan



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

longitudinal data. One approach for extending generalized linear models to longitudinal data is a class of regression models where the model for the mean response at each observation does not incorporate dependence on any random effects or previous responses. The model is known as a marginal model. An alternative approach for accounting for the within-subject association is via the introduction of random effects. The model is known as the generalized linear mixed effects model. Due to the interpretation of their regression coefficients, the former are often referred to as “population-average models,” and the latter are referred to as “subject-specific models” [1]. Assuming a situation where the target of inference is the population-level summary described as a basis for comparison between treatment conditions in ICH E9 (R1), “Addendum on Estimands and Sensitivity Analysis in Clinical Trials” to the guideline on Statistical Principles for Clinical Trials [2], we focus on marginal models here. For estimation of the marginal model parameters, assumptions about the joint distribution of the responses are not necessary. The avoidance of distributional assumptions leads to a method of estimation known as the generalized estimating equation (GEE) [3]. An appealing property of the GEE estimator is that it is a consistent estimator even if the assumed model for the covariance among the repeated measures is not correctly specified. In addition, we can obtain valid variance estimate for regression coefficients by utilizing the empirical or sandwich estimator. The sandwich estimator possesses a notable property as it demonstrates robustness by providing valid variance estimate even when the working covariance among the repeated measures is not correct. The GEE requires only that the model for the mean response, the link function and the variance of each response be correct.

For binary response data, recent studies have investigated the small-sample property of GEE estimates of the regression parameter and the bias-reduced estimates obtained by applying Firth’s adjustment for the likelihood score equation [4] to the GEE as if generalized estimating functions were likelihood score functions [5–8]. Paul and Zhang [5] found that GEE estimates yielded biased estimates of the regression parameters when the sample size was small and that bias-reduced estimates improved bias and mean squared error (MSE) by simulation studies. Mondol and Rahman [6] showed that bias-reduced GEE (also referred to as penalized GEE) achieved convergence and provided finite estimates in the presence of separation. Geroldinger et al. [7] reported that the bias-reduced estimation improved convergence compared to the standard GEE with a similar or even better performance in terms of the accuracy of estimates. Goshio et al. [8] reviewed bias-reduced GEEs and modified covariance

estimators and evaluated their performance in sparse binary data from small-sample longitudinal studies. In a study investigating marginal models in small samples, Bie et al. [9] reported that both GEE and marginalized multilevel models exhibit small-sample bias when correct correlation structure is adopted. Greenland [10] states that “the potential for small-sample bias in results from asymptotic procedures needs to be checked more routinely than is current practice.”

The proportional odds model is the commonly used model for relating ordinal outcome to covariates. The model is also referred to as the proportional odds version of the cumulative logit model [11]. The proportional odds model assumes that the association of each covariate with the outcome is represented by a single odds ratio. The purpose of this paper is to obtain the bias-reduced GEE (BR-GEE) estimates of the proportional odds model for longitudinal ordinal data and to investigate the small-sample properties of both the GEE and the BR-GEE through simulation. Kosmidis and Firth [12] derived general expressions for the adjusted score equations for the general class of multivariate models. Using the adjusted score equations exploited in Kosmidis and Firth [12] for the multivariate generalized linear model, Kosmidis [13] derived the adjusted score equation for the proportional odds model. We apply the adjusted likelihood score equation for the proportional odds model to the GEE to obtain the BR-GEE estimate as a solution to an adjusted estimating equation, which is induced by adding the adjustment function.

“Methods” section provides a brief summary of the BR-GEE estimates for longitudinal ordinal data. In “Simulation study” section, the results of the simulation study are presented and the small-sample properties of the GEE and the BR-GEE are investigated. In “Example” section, we apply the BR-GEE to the data of postoperative pain after laparoscopic cholecystectomy illustrated in Lumley [14]. A discussion of our findings follows in “Discussion” section.

## Methods

### Notation and GEE for the proportional odds model

Suppose a study includes  $N$  subjects with  $n_i$  repeated observations of the  $K$  multinomial ordered categories for the  $i$ th subject,  $i = 1, \dots, N$ . For simplicity, we assume equal repetition,  $n_i = n$ . Let  $Y_{it}$  represent responses as a vector  $Y_{it} = (Y_{it1}, \dots, Y_{itq})^T$  of  $q = K - 1$  dummy variables, where  $Y_{itk} = 1$  for the observation at occasion  $t$  ( $t = 1, \dots, n$ ) of subject  $i$  ( $i = 1, \dots, N$ ) falling in category  $k$  ( $k = 1, \dots, q$ ) and  $Y_{itk} = 0$  otherwise. The vector of marginal means or response probabilities of  $Y_{it}$  is  $\pi_{it} = (\pi_{it1}, \dots, \pi_{itq})^T$ . The proportional odds model links the cumulative probability

$\gamma_{itk} = \pi_{it1} + \dots + \pi_{itk}$  to a  $p$ -vector of covariates  $\mathbf{x}_{it}$  via the following relationship:

$$\log\left(\frac{\gamma_{itk}}{1-\gamma_{itk}}\right) = \beta_{0k} + \mathbf{x}_{it}^T \boldsymbol{\beta} \quad (i = 1, \dots, N; t = 1, \dots, n; k = 1, \dots, q),$$

$$E(Y_{itc} Y_{it'c'}) = \begin{cases} \gamma_{itct'/c'} & c = c' = 1 \\ \gamma_{itct'/c'} - \gamma_{itct'/(c'-1)} & c = 1, c' > 1 \\ \gamma_{itct'/c'} - \gamma_{it(c-1)t'/c'} & c > 1, c' = 1 \\ \gamma_{itct'/c'} - \gamma_{itct'/(c'-1)} - \gamma_{it(c-1)t'/(c'-1)} + \gamma_{it(c-1)t'/(c'-1)} & c > 1, c' > 1 \end{cases}$$

where  $\boldsymbol{\delta} = (\beta_{01}, \dots, \beta_{0q}, \boldsymbol{\beta}^T)^T$  is a  $(q+p)$ -vector of model parameters with  $\beta_{01} < \dots < \beta_{0q}$  and  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ . The linear term  $\eta_{itk} = \beta_{0k} + \mathbf{x}_{it}^T \boldsymbol{\beta}$  is determined by  $\eta_{itk} = \sum_{r=1}^{p+q} \delta_r z_{itkr}$ , where  $z_{itkr}$  is the  $(k, r)$ th element of the  $q \times (q+p)$  matrix,

$$\mathbf{Z}_{it} = \begin{pmatrix} 1 & 0 & \dots & 0 & \mathbf{x}_{it}^T \\ 0 & 1 & \dots & 0 & \mathbf{x}_{it}^T \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & \mathbf{x}_{it}^T \end{pmatrix} \quad (i = 1, \dots, N; t = 1, \dots, n).$$

Let  $\mathbf{Y}_i = (\mathbf{Y}_{i1}^T, \dots, \mathbf{Y}_{in}^T)^T$  and  $\boldsymbol{\pi}_i = (\boldsymbol{\pi}_{i1}^T, \dots, \boldsymbol{\pi}_{in}^T)^T$  denote responses and response probabilities, respectively, for subject  $i$ . Then, a GEE for consistent estimation of  $\boldsymbol{\delta}$  is given in the form,

$$\mathbf{U}(\boldsymbol{\delta}, \alpha) = \sum_{i=1}^N \mathbf{D}_i^T \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\pi}_i) = \mathbf{0},$$

where  $\mathbf{D}_i = D(\boldsymbol{\pi}_i; \boldsymbol{\delta})$  is the  $qn \times p$  Jacobian of  $\boldsymbol{\pi}_i$  with respect to  $\boldsymbol{\delta}$  and  $\mathbf{V}_i = \mathbf{V}_i(\boldsymbol{\delta}, \alpha)$  is a ‘working’ covariance matrix indexed by  $\boldsymbol{\delta}$  and an association parameter  $\alpha$ . Further details on the GEE for ordinal data can be found in Agresti [11], Heagerty and Zeger [15], Fahrmeir and Pritscher [16], Lipsitz et al. [17], Liang et al. [18], and Williamson et al. [19].

Let  $\mathbf{V}_{itt} = \text{Var}(\mathbf{Y}_{it}) = \text{diag}(\boldsymbol{\pi}_{it}) - \boldsymbol{\pi}_{it} \boldsymbol{\pi}_{it}^T$  denote the  $q \times q$  multinomial covariance matrix for  $\mathbf{Y}_{it}$ . For  $t \neq t'$ ,  $\mathbf{V}_{itt'} = \text{Cov}(\mathbf{Y}_{it}, \mathbf{Y}_{it'})$  is the covariance matrix between two different occasions in the same subject. The simplest model for the working covariance matrix is the independent model,  $\mathbf{V}_{itt'} = \mathbf{0}$ . Lumley [14] described working covariance structures  $\mathbf{V}_{itt'}$  and computational methods for ordinal data based on the global odds ratio that allow the GEE to be used for smaller sets of ordinal data and with less effort expended on modeling associations. For each pair of categories  $c$  and  $c'$  of two ordinal responses  $R_{it}$  and  $R_{it'}$  in the same subject, the global odds ratio is defined as

$$\Psi_{itct'/c'} = \frac{\Pr(R_{it} \leq c, R_{it'} \leq c') \Pr(R_{it} > c, R_{it'} > c')}{\Pr(R_{it} > c, R_{it'} \leq c') \Pr(R_{it} \leq c, R_{it'} > c')}, \quad c, c' = 1, \dots, q.$$

The covariance of  $Y_{itc}$  and  $Y_{it'c'}$  is given by  $\text{cov}(Y_{itc}, Y_{it'c'}) = E(Y_{itc} Y_{it'c'}) - E(Y_{itc}) E(Y_{it'c'})$ , where the second term is the product of  $\pi_{itc}$  and  $\pi_{it'c'}$ . The first term  $E(Y_{itc} Y_{it'c'})$  can be expressed by the joint cumulative probabilities  $\gamma_{itct'/c'} = \Pr(R_{it} \leq c, R_{it'} \leq c')$ :

The joint cumulative probabilities  $\gamma_{itct'/c'}$  can be expressed in terms of global odds ratio and marginal cumulative probabilities:

$$\gamma_{itct'/c'} = \frac{\kappa - \sqrt{\kappa^2 - 4\Psi_{itct'/c'}(\Psi_{itct'/c'} - 1)\gamma_{itc}\gamma_{it'c'}}}{2(\Psi_{itct'/c'} - 1)}, \quad (1)$$

where  $\kappa = 1 + (\gamma_{itc} + \gamma_{it'c'}) (\Psi_{itct'/c'} - 1)$ . A simple estimate of the crude global odds ratio is calculated by the weighted mean of the log odds ratios with weights inversely proportional to the variances for every possible pair of the two time points  $(t, t')$ , where the odds ratios are computed based on the  $2 \times 2$  tables obtained from the  $q^2$  ways of collapsing the row and column classifications into dichotomies for a  $K \times K$  table of the variables  $R_{it}$  and  $R_{it'}$  over all subjects. Hines [20] compared the various models for  $\mathbf{V}_{itt'}$  both in terms of efficiency and computational stability. She showed that the performance of Lumley’s crude global odds ratio method is satisfactory. Therefore, our approach is based on the working covariance structure of the independence or exchangeable model, where the crude global odds ratio is constant for all pairs of times.

#### Bias-reduced GEE for the proportional odds model

Firth [4] showed that the solution of the following adjusted score equation results in an estimator that is free from the first-order term in the asymptotic expansion of the bias in the MLE of the regression parameter:

$$\mathbf{U}^*(\boldsymbol{\theta}) = \mathbf{U}(\boldsymbol{\theta}) + \mathbf{A}(\boldsymbol{\theta}) = \mathbf{0},$$

with  $\mathbf{A}(\boldsymbol{\theta}) = -\mathbf{I}(\boldsymbol{\theta}) \mathbf{b}(\boldsymbol{\theta})$ , where  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^T$  is a  $p$ -vector of regression parameters from the generalized linear model,  $\mathbf{U}(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} l(\boldsymbol{\theta}) = \mathbf{0}$  is the standard score equation based on the log-likelihood function  $l(\boldsymbol{\theta})$ ,  $\mathbf{I}(\boldsymbol{\theta})$  is the expected information matrix for  $\boldsymbol{\theta}$ , and  $\mathbf{b}(\boldsymbol{\theta})$  is the first term in the asymptotic expansion of the bias of the MLE.

Kosmidis and Firth [12] derived general expressions for the adjusted score functions for the general class of multivariate models, which include multivariate generalized linear models. Using the adjustment functions exploited in Kosmidis and Firth [12] for the score functions of the

multivariate generalized linear model, Kosmidis [13] formulated the adjustment function of the score function of the proportional odds model. Considering that the GEE is equivalent to the generalized linear model score function under the identity working covariance structure, we treat  $\mathbf{U}(\boldsymbol{\delta}, \alpha)$  as if it were a likelihood score function and apply the adjustment function in the score function to the GEE. Then, the adjustment function for the generalized estimating function for the proportional odds model is given by

$$A_r(\boldsymbol{\delta}, \alpha) = \frac{1}{2} \sum_{i=1}^N \sum_{s=1}^{qn} \text{tr} \left[ \mathbf{Z}_i \mathbf{F}^{-1} \mathbf{Z}_i^T \left\{ \left( D(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i) \mathbf{V}_i^{-1} \right)_s \otimes \mathbf{I}_{qn} \right\} D^2(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i) \right] z_{isr} \quad (r = 1, \dots, q + p),$$

where  $\mathbf{Z}_i = (\mathbf{Z}_{i1}^T, \dots, \mathbf{Z}_{in}^T)^T$ ,  $\boldsymbol{\eta}_i = (\boldsymbol{\eta}_{i1}^T, \dots, \boldsymbol{\eta}_{in}^T)^T$  with elements of  $\boldsymbol{\eta}_{it} = (\eta_{it1}, \dots, \eta_{itq})^T$ ,  $D(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i)$  is the  $qn \times qn$  Jacobian of  $\boldsymbol{\pi}_i$  with respect to  $\boldsymbol{\eta}_i$ ,  $\mathbf{F} = \sum_{i=1}^N (D(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i) \mathbf{Z}_i^T \mathbf{V}_i^{-1} D(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i) \mathbf{Z}_i)$ ,  $D^2(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i)$  is the  $(qn)^2 \times qn$  matrix with the  $s$ th block as the Hessian of  $\pi_{is}$  with respect to  $\boldsymbol{\eta}_i$  ( $s = 1, \dots, qn$ ),  $\mathbf{I}_{qn}$  is the  $qn \times qn$  identity matrix,  $z_{isr}$  is the  $(s, r)$ th element of  $\mathbf{Z}_i$ ,  $\mathbf{V}_i$  is the covariance matrix of the vector  $\mathbf{Y}_i$ , and  $(D(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i) \mathbf{V}_i^{-1})_s$  is the  $s$ th row of  $D(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i) \mathbf{V}_i^{-1}$ . The matrix  $D(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i)$  is the block diagonal matrix whose  $t$ th diagonal element is

$$\Delta_{it} = \begin{pmatrix} g_{it1} & 0 & \cdots & 0 & 0 \\ -g_{it1} & g_{it2} & \cdots & 0 & 0 \\ 0 & -g_{it2} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & g_{it,q-1} & 0 \\ 0 & 0 & \cdots & -g_{it,q-1} & g_{itq} \end{pmatrix} \quad (i = 1, \dots, N; t = 1, \dots, n),$$

where  $g_{itk} = g(\eta_{itk})$  with  $g(\eta) = dG(\eta)/d\eta$  and  $G(\eta) = \exp(\eta)/(1 + \exp(\eta))$ ; hence,  $g_{itk} = G(\eta_{itk})(1 - G(\eta_{itk})) = \gamma_{ikt}(1 - \gamma_{ikt})$ .

$$\boldsymbol{\Sigma}_{GEE} = \left( \sum_{i=1}^N \mathbf{D}_i^T \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1} \left( \sum_{i=1}^N \mathbf{D}_i^T \mathbf{V}_i^{-1} \text{Cov}(\mathbf{Y}_i) \mathbf{V}_i^{-1} \mathbf{D}_i \right) \left( \sum_{i=1}^N \mathbf{D}_i^T \mathbf{V}_i^{-1} \mathbf{D}_i \right)^{-1},$$

The matrix  $D^2(\boldsymbol{\pi}_i; \boldsymbol{\eta}_i)$  is the  $(qn)^2 \times qn$  matrix with the  $(s, u)$ th element given by:

$$c_{su} = \begin{cases} \gamma_{itk}(1 - \gamma_{itk})(1 - 2\gamma_{itk}), & s = (v - 1)(nq + 1) + 1, u = v \\ -\gamma_{itk}(1 - \gamma_{itk})(1 - 2\gamma_{itk}) & s = w(nq + 1) + 1, u = w \\ 0, & \text{otherwise} \end{cases}$$

where  $v = (t - 1)q + k$  ( $t = 1, \dots, n; k = 1, \dots, q$ ) and  $w = (t - 1)q + k$  ( $t = 1, \dots, n; k = 1, \dots, q - 1$ ).

The  $r$ th element of the bias-reduced generalized estimating function is  $U_r^*(\boldsymbol{\delta}, \alpha) = U_r(\boldsymbol{\delta}, \alpha) + A_r(\boldsymbol{\delta}, \alpha)$ . The

BR-GEE estimate  $\tilde{\boldsymbol{\delta}}$  is such that  $U_r^*(\tilde{\boldsymbol{\delta}}, \tilde{\alpha}) = 0$  for every  $r = 1, \dots, q + p$ , where  $\tilde{\alpha}$  is the estimate of global odds ratio.  $\tilde{\boldsymbol{\delta}}$  can be estimated using an iterative method described below:

Step 1: Choose an initial value  $\tilde{\boldsymbol{\delta}}^0$  of  $\boldsymbol{\delta}$ .

Step 2: If the working covariance structure is exchangeable, we calculate the log (crude) global odds ratio  $\tilde{\alpha}$  [14].

Step 3: Given  $\tilde{\boldsymbol{\delta}}^{(t)}$  at the  $t$ th iteration and  $\tilde{\alpha}$  (unnecessary if the working covariance structure is independent),

the covariance matrix is estimated from the global odds ratio and the predicted cumulative probabilities [14].

Step 4: Given the working covariance matrix  $\mathbf{V}_i^{-1}(\tilde{\boldsymbol{\delta}}^{(t)}, \tilde{\alpha})$ , the current estimate  $\tilde{\boldsymbol{\delta}}^{(t)}$  is updated according to the adjusted GEE using the Newton–Raphson method given by

$$\tilde{\boldsymbol{\delta}}^{(t+1)} = \tilde{\boldsymbol{\delta}}^{(t)} + \mathbf{F}(\boldsymbol{\delta}, \tilde{\alpha})^{-1} \Big|_{\boldsymbol{\delta}=\tilde{\boldsymbol{\delta}}^{(t)}} \mathbf{U}^*(\boldsymbol{\delta}, \tilde{\alpha}) \Big|_{\boldsymbol{\delta}=\tilde{\boldsymbol{\delta}}^{(t)}},$$

with

$$\mathbf{F}(\boldsymbol{\delta}, \alpha) = \sum_{i=1}^N \mathbf{D}_i^T \mathbf{V}_i^{-1} \mathbf{D}_i.$$

Step 5: Iterate steps 3 and 4 until a desired convergence criterion is satisfied (for example,  $\max |\tilde{\boldsymbol{\delta}}^{(t+1)} - \tilde{\boldsymbol{\delta}}^{(t)}| \leq 0.0001$ ). At convergence, the estimate of  $\boldsymbol{\delta}$  is denoted by  $\tilde{\boldsymbol{\delta}}$  and referred to as the BR-GEE estimate.

The consistent estimate of the covariance matrix of the GEE regression parameter estimates is given by the sandwich estimator of Liang and Zeger [3],

where  $\hat{\mathbf{r}}_i \hat{\mathbf{r}}_i^T = (\mathbf{Y}_i - \hat{\boldsymbol{\pi}}_i)(\mathbf{Y}_i - \hat{\boldsymbol{\pi}}_i)^T$  is used to estimate  $\text{Cov}(\mathbf{Y}_i)$ .

In this manuscript, we calculate the covariance matrix for the BR-GEE derived from the estimating function  $\mathbf{U}^*(\boldsymbol{\delta}, \alpha) = \mathbf{U}(\boldsymbol{\delta}, \alpha) + \mathbf{A}(\boldsymbol{\delta}, \alpha)$ ,

$$\Sigma_{BR-GEE} = \left( \sum_{i=1}^N D_i^T V_i^{-1} D_i + A A^T \right)^{-1} M \left( \sum_{i=1}^N D_i^T V_i^{-1} D_i + A A^T \right)^{-1},$$

with

$$M = \left( \sum_{i=1}^N D_i^T V_i^{-1} \hat{r}_i \hat{r}_i^T V_i^{-1} D_i + A A^T \right).$$

Mancl and DeRouen [21] described that the bias-corrected covariance estimator used with the critical value based on the F-distribution instead of chi-square produced proper test sizes. Therefore, in the following, we used the critical value of the t-distribution with the number of the sample size minus the number of coefficients in the regression model as the degree of freedom to compute the confidence interval and the statistical significance.

When the sample size is small, the sandwich covariance matrix is expected to underestimate the covariance matrix. To reduce the small-sample bias, several modified variance estimators have been developed [22]. In this manuscript, in addition to  $\Sigma_{GEE}$  and  $\Sigma_{BR-GEE}$ , we used the bias-corrected covariance estimates of Mancl and DeRouen [21] obtained by substituting the GEE and BR-GEE estimates into the following equation in order to calculate the 95% confidence interval,

$$\Sigma_{MD} = \left( \sum_{i=1}^N D_i^T V_i^{-1} D_i \right)^{-1} M_{MD} \left( \sum_{i=1}^N D_i^T V_i^{-1} D_i \right)^{-1}.$$

$$\text{logit}(\gamma_{itk}) = \beta_{0k} + \beta_1 \text{Trt}_i + \sum_{s=1}^3 \{ \beta_{s+1} I(\text{Time}_{it} = s) + \beta_{s+4} \text{Trt}_i \times I(\text{Time}_{it} = s) \} \quad (k = 1, 2), \quad (2)$$

with

$$M_{MD} = \left( \sum_{i=1}^N D_i^T V_i^{-1} (I_{qn} - H_{ii}) \hat{r}_i \hat{r}_i^T (I_{qn} - H_{ii})^T V_i^{-1} D_i \right),$$

where  $I_{qn}$  is the  $qn \times qn$  identity matrix and  $H_{ii}$  is the matrix calculated by

$$H_{ii} = D_i \left( \sum_{i=1}^N D_i^T V_i^{-1} D_i \right)^{-1} D_i^T V_i^{-1}.$$

### Simulation study

In this section, we conducted a limited simulation study to investigate the small-sample properties of both the standard GEE (GEE) and the proposed bias-reduced GEE

(BR-GEE) estimation of the regression parameters in a marginal model for correlated ordinal data. We assumed a randomized clinical trial with a parallel group design, where each subject was assigned to either the treatment group or the control group. Balanced correlated ordinal data with  $K = 3$  of the total number of response categories and with observations over  $n = 4$  time occasions were generated using the algorithm given in Ibrahim et al. [23]. They used the Goodman and Kruskal  $\Gamma$  coefficient as a measurement of the association for the responses of each subject. We modified the algorithm to specify the association of the within-subject observation by using the global odds ratio. Specifically, we modified the macro developed by Ibrahim et al. [23] to calculate the joint probabilities of responses between two time points, which are used to generate correlated ordinal data, by substituting the global odds ratio and marginal cumulative probability based on the marginal model into expression (1). For the exchangeable correlation, we used a common  $\exp(\alpha)$  as global odds ratio for all pairs of time points, while for AR-type correlation, we used  $\exp(\alpha/\tau)$  as global odds ratio based on the time interval  $\tau$  for each pair of time points. We used a proportional odds marginal model both to simulate and to fit the data. We considered the following proportional odds model for the  $i$  th subject measured at the  $t$  th occasion ( $i = 1, \dots, N; t = 1, \dots, 4$ ):

where  $\gamma_{itk}$  is the cumulative probability,  $\text{Trt}_i$  is a treatment assignment variable coded as 1 for the first half of the subjects or 0 for the second half of the subjects, and  $\text{Time}_{it}$  is an occasion coded as 1, 2, 3 and 4. We conducted simulations for four different sample sizes:  $N = 20, 30, 40, 50$ . We let the  $(q + p)$ -vector parameter set used with regression model (2) be  $\delta = (\beta_{01}, \beta_{02}, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7)^T = (-0.1, 1, 1.2, -0.9, -0.6, -0.3, -0.3, -0.2, -0.1)^T$  to evaluate the performance under the alternative hypothesis (Scenario 1). In addition, we let  $\delta = (1.1, 2.2, 0, -2.1, -1.4, -0.7, 0, 0, 0)^T$  to evaluate the empirical type I error rate (Scenario 2). For this set of simulations, the marginal response probabilities of the treatment group at the last time point are (0.75, 0.15, 0.10). The covariance structure used to generate data was an exchangeable structure with the log global odds ratio of  $\alpha$  as 1, 1.5, 2 and an autoregressive-type (AR-type)

structure with the log global odds ratio of  $\alpha/|t - t'| (t \neq t')$  for each  $\alpha$  as 1, 1.5, 2. For each parameter configuration, 5,000 simulation replications were performed.

For each scenario, we fit the correct marginal mean model, and the covariate coefficients were estimated by the GEE and the BR-GEE with the independent and exchangeable in terms of the crude global odds ratio suggested in Lumley [14] as the working covariance structure. The algorithm convergence criterion was less than or equal to 0.0001, and the maximum iterations allowed were set to 50. The standard errors of the GEE and the BR-GEE estimated regression coefficients were estimated by substituting each parameter estimate into  $\Sigma_{GEE}$  and  $\Sigma_{BR-GEE}$ , respectively. We defined a model fit as having a convergence problem if the algorithm convergence criterion was not met or if the maximum of the absolute value of the estimated regression parameters was above 10. There were a small number of datasets where the diagonal elements of the sandwich variance estimate were negative when there was a convergence problem. At that time, we replaced it with the model-based variance to evaluate the performance measures. We focused on the results of the coefficient  $\beta_1$ , which represents the treatment effect at the last time point. In this simulation study, we defined the bias as  $\text{bias}(\hat{\beta}_1) = \hat{\beta}_1 - \beta_1$ , where  $\hat{\beta}_1 = \sum_{l=1}^{L_s} \hat{\beta}_{1l} / L_s$ ,  $\hat{\beta}_{1l}$  is the estimate of  $\beta_1$  for the  $l$ th replication, and  $L_s$  is the number of runs in which the regression parameter was estimated. The root mean square error (RMSE) was defined by  $\text{RMSE}(\hat{\beta}_1) = \sqrt{\sum_{l=1}^{L_s} (\hat{\beta}_{1l} - \beta_1)^2 / L_s}$ . Using the standard error estimate, we constructed 95% confidence intervals by multiplying the standard error with the 2.5th percentile and 97.5th percentile of the t-distribution. Then, we defined the coverage probability of 95% confidence intervals for the parameter  $\beta_1$  to be the proportion of the number of 95% confidence intervals that contain the true parameter value to the total number of runs where the regression parameter was estimated. Furthermore, the empirical type I error rate was defined as the proportion of times  $|\hat{\beta}_{1l} / \text{SE}(\hat{\beta}_{1l})| \geq t_{0.975}$  for the null hypothesis, where  $t_{0.975}$  is the 97.5th percentile of the t-distribution.

In addition, the bias, RMSE and coverage for the specified scenarios with  $(K, n) = (3, 6), (4, 4), (4, 6)$  were investigated. And, for each scenario, coverage probability of the 95% confidence interval based on the Mancl and DeRouen's bias-corrected covariance estimates were calculated in addition to that based on  $\Sigma_{GEE}$  and  $\Sigma_{BR-GEE}$ .

### Results for the convergence

First, we report the percentage of simulation replications in which there was evidence of quasicomplete separation as well as the percentage of simulation sets

**Table 1** Example of quasicomplete separation due to treatment (Trt) against outcome  $Y$  ( $N = 40$ )

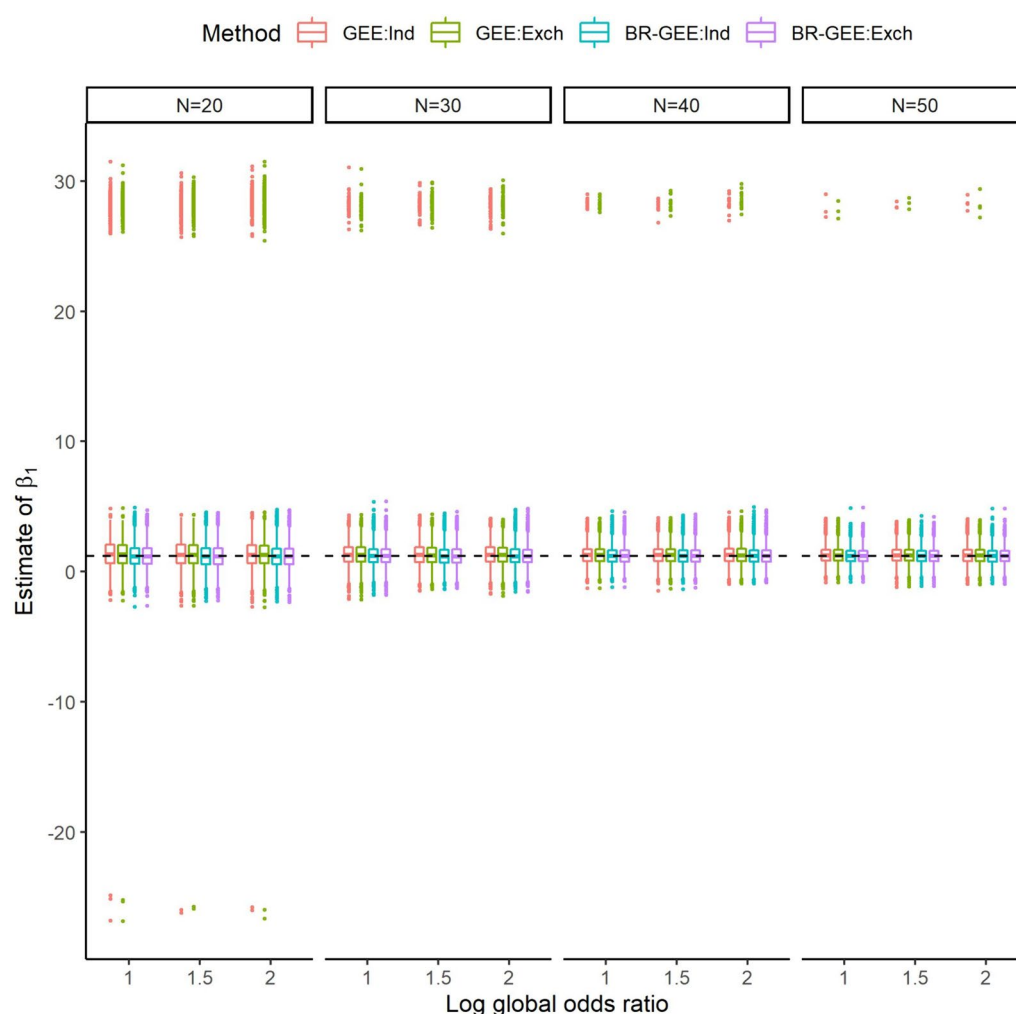
|     |   | Quasicomplete separation |   |    |
|-----|---|--------------------------|---|----|
|     |   | $Y$                      |   |    |
|     |   | 1                        | 2 | 3  |
| Trt | 1 | 20                       | 0 | 0  |
|     | 0 | 5                        | 5 | 10 |

where the model had convergence problems for Scenario 1 and Scenario 2 (see Table S1 for the results with true exchangeable covariance structure (Scenario 1), Table S2 for the results with true exchangeable covariance structure (Scenario 2), Table S3 for the results with true AR-type covariance structure (Scenario 1), and Table S4 for the results with true AR-type covariance structure (Scenario 2) in Additional file). We defined that the simulation datasets had evidence of quasicomplete separation when all categories but one level of 1 or  $K$  of the outcome for either treatment group were zero at a time point, such as Table 1. The model fit by the GEE had the convergence problem in the presence of quasicomplete separation. For instance, the percentage of datasets where GEE estimation had a convergence problem fell within the range 7.30–8.06 when  $N = 20$ , 1.56–1.66 when  $N = 30$ , 0.28–0.40 when  $N = 40$ , and 0.06–0.08 when  $N = 50$  for Scenario 1 with true exchangeable covariance structure. In contrast, it was 0.00% for all settings for the BR-GEE. The performance of the estimates of  $\beta_1$  was similar for both true covariance structures. Therefore, here, we present the results of the estimates of  $\beta_1$  in the true exchangeable covariance structure only. See Additional file for the results in the AR-type covariance structure.

### Results for the performance under the alternative hypothesis (Scenario 1)

In this section, we first report the distribution of the estimates. Then, we summarize the results for the simulation in terms of bias, RMSE, and coverage probability of the 95% confidence interval.

Figure 1 shows the box and whisker plot for estimates of  $\beta_1$  with an exchangeable (correctly specified) and an independent (misspecified) working covariance for the true exchangeable covariance structure. The distribution of the GEE estimates had substantially large outliers for datasets with evidence of quasicomplete separation. On the other hand, the BR-GEE did not have such outliers, regardless of the value of the global odds ratio and the working covariance structure for each sample size ( $N \leq 50$ ). See Fig. S1 for the results in the AR-type covariance structure.



**Fig. 1** Box and whisker plot for estimates of  $\beta_1$  with true exchangeable covariance structure (Scenario 1). GEE:Ind, generalized estimating equation with working independent covariance structure, GEE:Exch, generalized estimating equation with working exchangeable covariance structure, BR-GEE:Ind, bias-reduced generalized estimating equation with working independent covariance structure, BR-GEE:Exch, generalized estimating equation with working exchangeable covariance structure

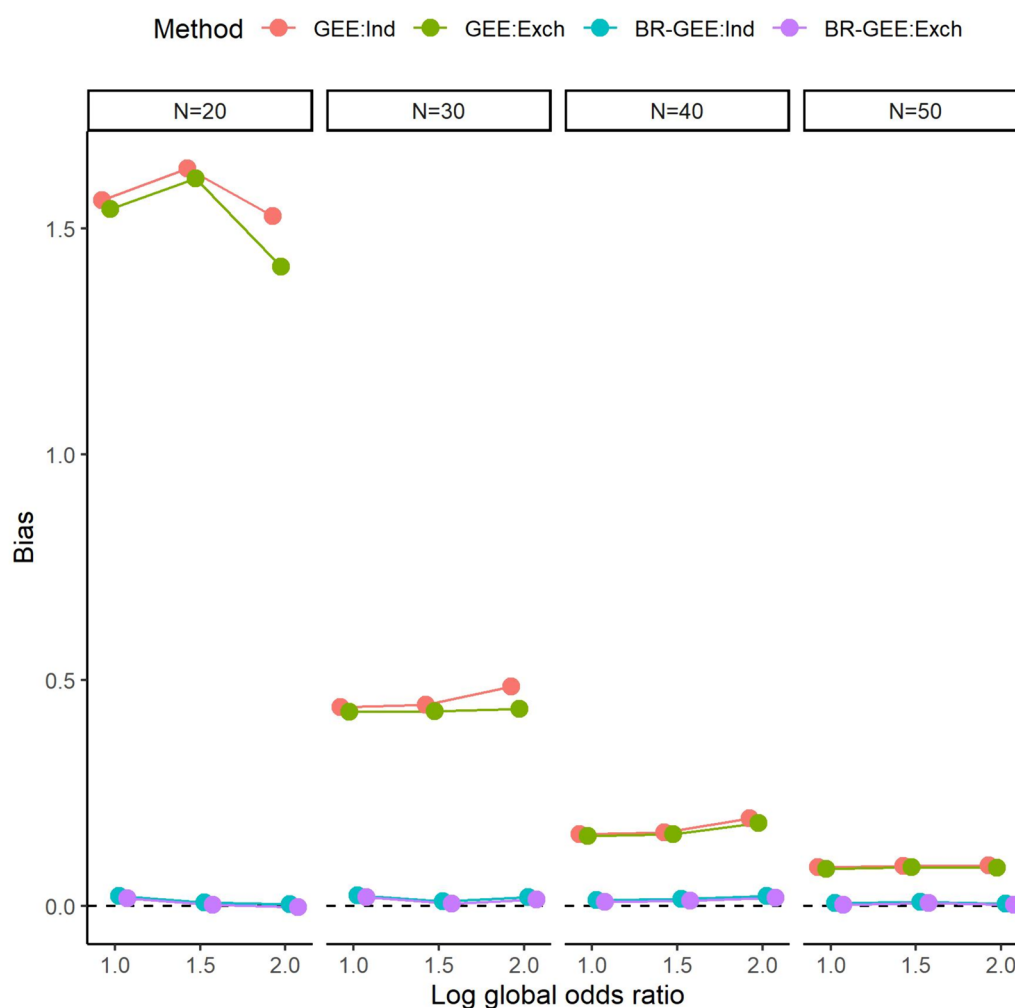
Figure 2 shows the bias of the estimates of  $\beta_1$ . Although it decreased as the sample size increased, bias of the GEE estimates still existed when  $N = 50$ . On the other hand, the bias of the BR-GEE estimates was close to zero regardless of the value of the global odds ratio and the working covariance structure for each sample size ( $N \leq 50$ ). See Fig. S2 for the results in the AR-type covariance structure.

Figure 3 shows the RMSE of the estimates of  $\beta_1$ . The RMSE of the BR-GEE estimates was smaller than that of the GEE estimates regardless of the value of the global odds ratio and the working covariance structure for each sample size ( $N \leq 50$ ). See Fig. S3 for the results in the AR-type covariance structure.

Figure 4 shows the coverage probabilities of the 95% confidence interval of  $\beta_1$ . For  $N=20$ , the coverage

probability of the 95% confidence interval based on the GEE was below the nominal level of 95%; on the other hand, that based on the BR-GEE was close to 95%. In general, as the number of subjects increased ( $N \geq 30$ ), the coverage probabilities of both methods became similar and slightly above 95%. See Fig. S4 for the results in the AR-type covariance structure. When the 95% confidence interval based on the BR-GEE was estimated by substituting the parameter estimate into  $\Sigma_{GEE}$  instead of  $\Sigma_{BR-GEE}$ , undercoverage of the 95% confidence interval based on the BR-GEE was observed for  $N=20$  but not for  $N \geq 30$ . See Fig. S6 for the results in the exchangeable covariance structure, and Fig. S7 for those in the AR-type covariance structure.

The bias and RMSE for the scenarios with  $(K, n) = (3, 6), (4, 4), (4, 6)$  were similar to the results when

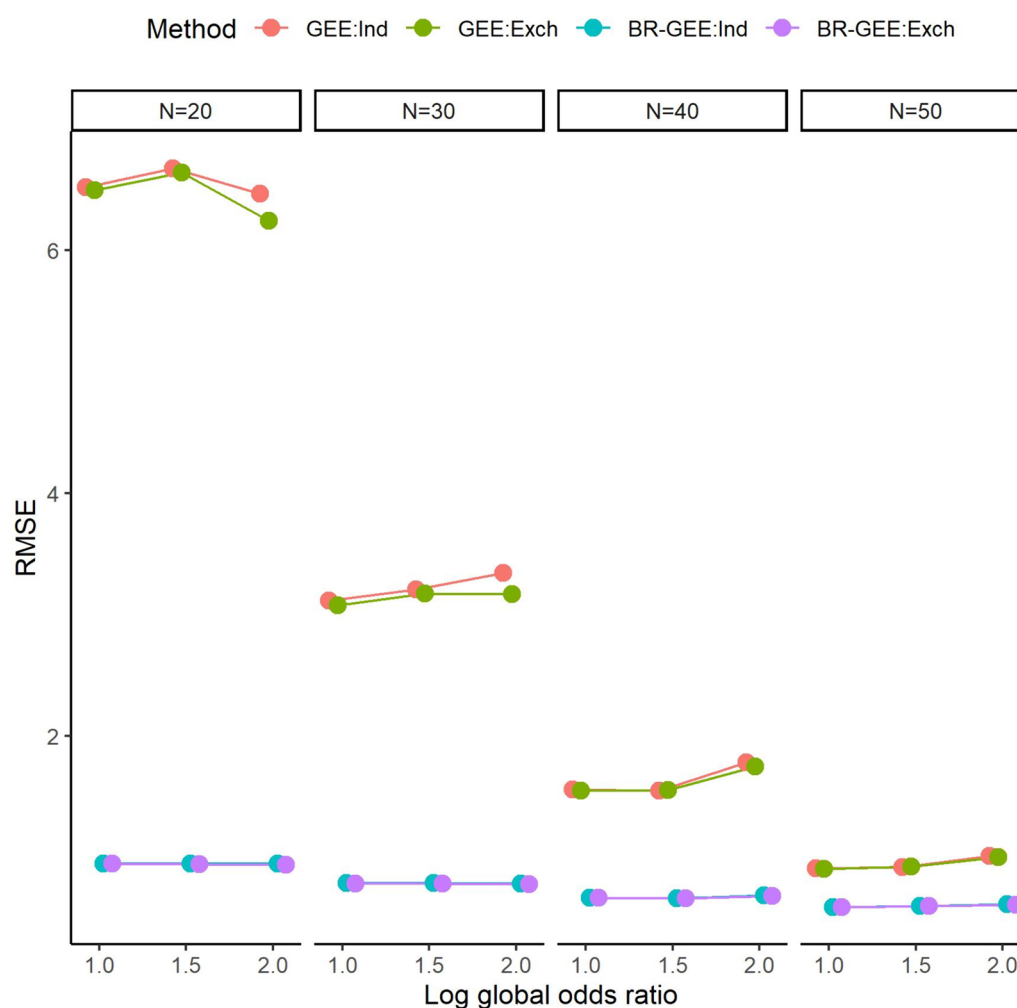


**Fig. 2** Bias associated with estimates of  $\beta_1$  with true exchangeable covariance structure (Scenario 1). GEE:Ind, generalized estimating equation with working independent covariance structure, GEE:Exch, generalized estimating equation with working exchangeable covariance structure, BR-GEE:Ind, bias-reduced generalized estimating equation with working independent covariance structure, BR-GEE:Exch, generalized estimating equation with working exchangeable covariance structure

$K = 3$  and  $n = 4$ . The differences in coverage probabilities between GEE and BR-GEE exhibited similar characteristics for  $(K, n) = (3, 4)$  and  $(3, 6)$ . Specifically, when  $N = 20$ , the coverage probabilities based on the GEE fell below the nominal level while those based on the BR-GEE exceeded it. However, when  $K = 4$ , such differences in coverage probabilities between GEE and BR-GEE were not observed. See Fig. S10 to Fig. S27 in Additional file. The coverage probability of the 95% confidence interval based on the Mancl and DeRouen's bias-corrected covariance estimates calculated by substituting the BR-GEE estimates generally remained above the nominal confidence level, similar to that of the 95% confidence interval based on  $\Sigma_{BR-GEE}$ . See Table. S5 to Table. S12 in Additional file.

### Results for the type I error rate under the null hypothesis (Scenario 2)

Figure 5 presents the empirical type I error rate of the t-test under the null hypothesis  $H_0 : \beta_1 = 0$  resulting from the use of the GEE or the BR-GEE. For  $N=20$ , although the type I error rate for the GEE was somewhat inflated, that of the BR-GEE was approximately 0.05. In general, as the number of subjects increased ( $N \geq 30$ ), the type I error rates of both methods became similar and slightly conservative. See Fig. S5 for the results in the AR-type covariance structure. For the empirical type I error rate of the t-test under the null hypothesis,  $H_0 : \beta_1 = 0$ , when the standard error for the BR-GEE as well as the GEE was estimated using the standard sandwich variance estimator, inflation of the type I error rate was observed



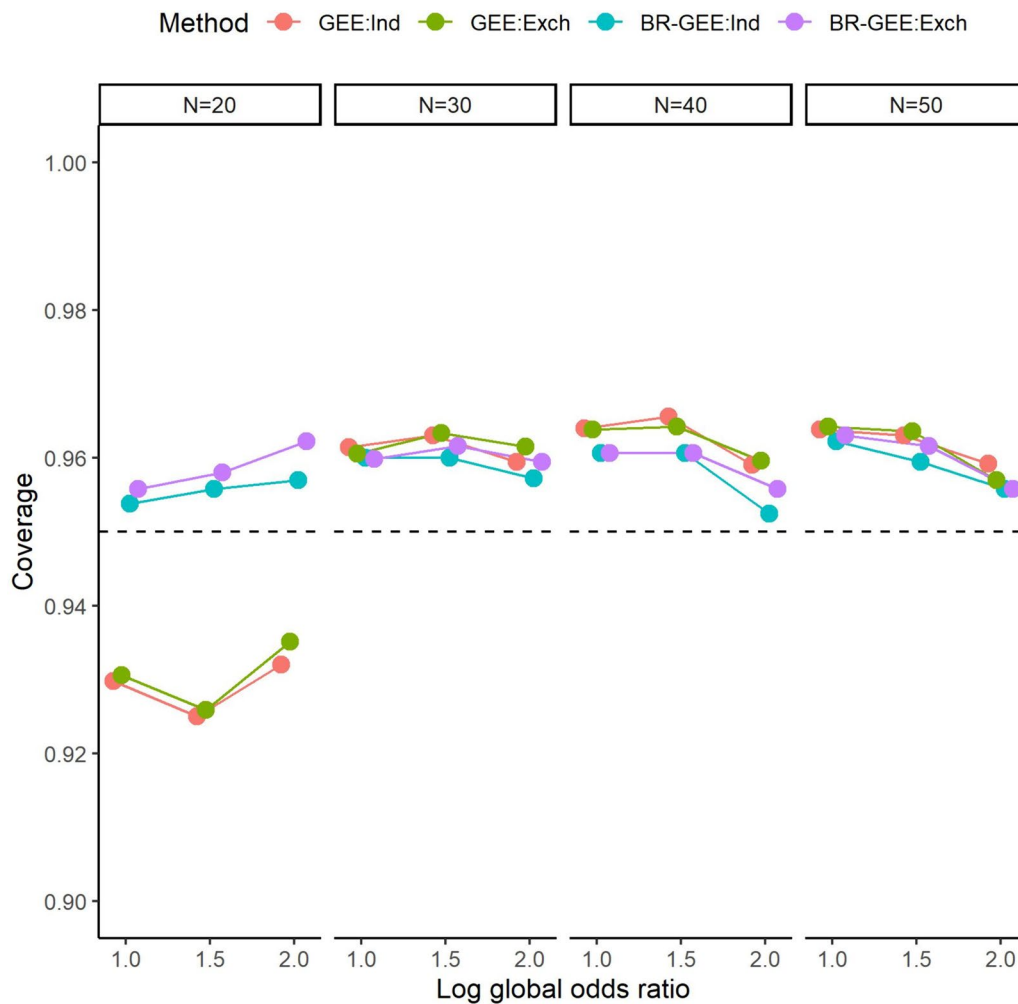
**Fig. 3** RMSE associated with estimates of  $\beta_1$  with true exchangeable covariance structure (Scenario 1). GEE:Ind, generalized estimating equation with working independent covariance structure, GEE:Exch, generalized estimating equation with working exchangeable covariance structure, BR-GEE:Ind, bias-reduced generalized estimating equation with working independent covariance structure, BR-GEE:Exch, generalized estimating equation with working exchangeable covariance structure

for  $N=20$  but not for  $N \geq 30$ . See Fig. S8 for the results in the exchangeable covariance structure, and Fig. S9 for those in the AR-type covariance structure.

### Example

In this section, we applied the proposed method to data from a randomized clinical trial of patients with postoperative pain after laparoscopic cholecystectomy. One of the aims of the study was to determine whether the use of an abdominal suction drain reduces shoulder tip pain after laparoscopic surgery. Patients rated their shoulder pain levels on a visual analog scale in the morning and afternoon of the first 3 days after the operation. The dataset from Lumley [14] contained 41 patients (22 patients

in the active treatment group and 19 patients in the control group) with 6 time points and 5 ordered pain score categories per patient. The pain in the control group peaked on the afternoon of day 2 [24]. Therefore, in this manuscript, we analyzed the data of 4 time points (Day 1 (am), Day 1 (pm), Day 2 (am), and Day 2 (pm)) of all 6 time points and estimated the treatment effect at Day 2 (pm). The time points for each morning and afternoon from Day 1 to Day 2 are designated as time=1 to time=4 in sequential order. To estimate the treatment effect at the last time point, we fit the marginal model of the proportional odds model with 9 covariates: 8 binary indicators (treatment, sex, time point, interaction of treatment and time point) and one continuous variable (age),

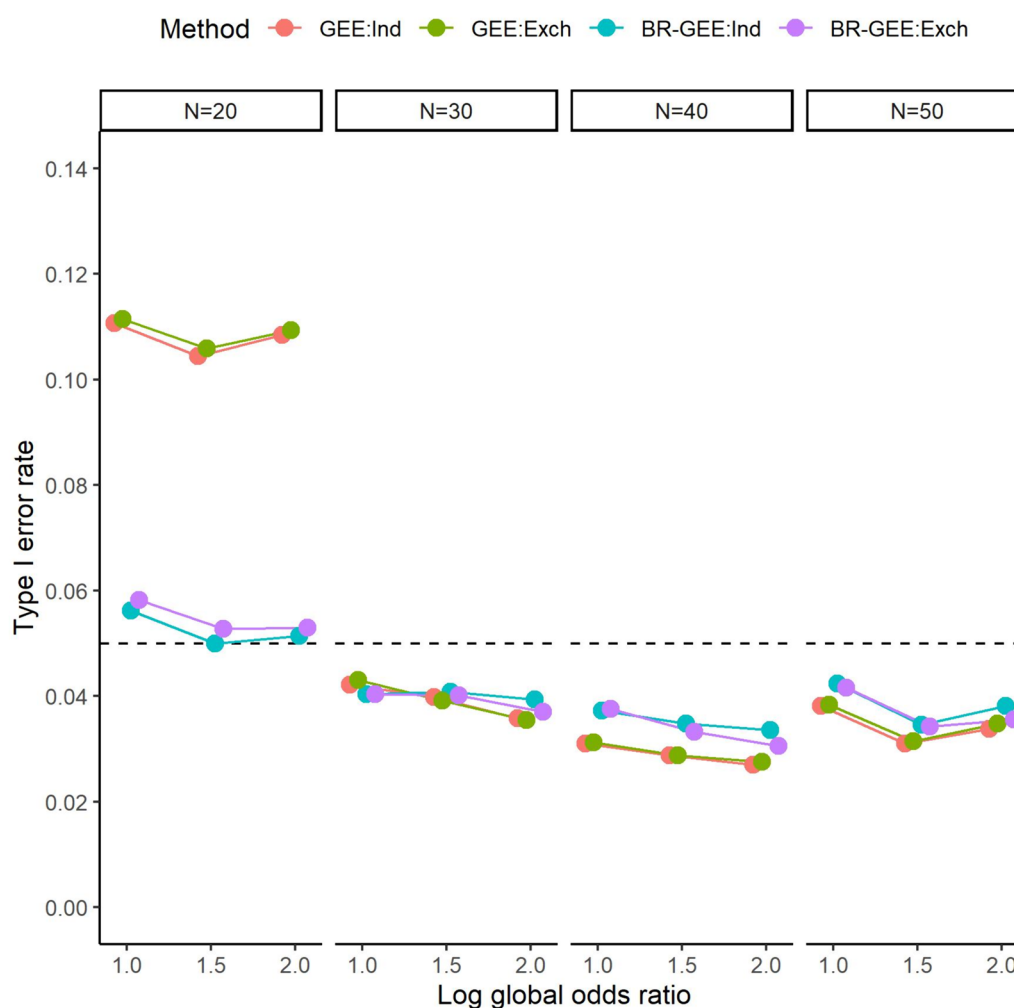


**Fig. 4** 95% confidence interval coverage for estimates of  $\beta_1$  with true exchangeable covariance structure (Scenario 1). GEE:Ind, generalized estimating equation with working independent covariance structure, GEE:Exch, generalized estimating equation with working exchangeable covariance structure, BR-GEE:Ind, bias-reduced generalized estimating equation with working independent covariance structure, BR-GEE:Exch, generalized estimating equation with working exchangeable covariance structure

$$\text{logit}(\gamma_{itk}) = \beta_{0k} + \beta_1 I(\text{Trt}_i = 1) + \beta_2 I(\text{Sex}_i = 1) + \beta_3 \text{Age}_i + \sum_{s=1}^3 \{ \beta_{s+3} I(\text{Time}_{it} = s) + \beta_{s+6} I(\text{Trt}_i = 1) \times I(\text{Time}_{it} = s) \} \quad (k = 1 \text{ to } 4),$$

where  $\text{Sex}_i = 1$  denotes that the subject is male,  $\text{Trt}_i = 1$  if the subject receives the active treatment, and 0 otherwise. For this model, we estimated the regression parameter by using the GEE and the BR-GEE, assuming an independent and exchangeable working covariance structure. In addition, we calculated the 95% confidence interval by using the t-distribution. In this model, the regression parameter  $\beta_1$  represents the treatment effect at the last time point ( $\text{Time} = 4$ ).

The odds ratio of favorable response comparing the active treatment group with the control group and the 95% confidence intervals by using the GEE with independent working covariance structure, the GEE with exchangeable working covariance structure, the BR-GEE with independent working covariance structure, and the BR-GEE with exchangeable working covariance structure were 14.0 (3.6, 54.5), 12.8 (3.5, 47.0), 12.1 (3.6, 41.1) and 10.6 (3.2, 34.8), respectively. Each of the odds ratios can be interpreted as a common odds ratio derived from



**Fig. 5** Type I error rate of t-test of  $H_0 : \beta_1 = 0$  with true exchangeable covariance structure (Scenario 2). GEE:Ind, generalized estimating equation with working independent covariance structure, GEE:Exch, generalized estimating equation with working exchangeable covariance structure, BR-GEE:Ind, bias-reduced generalized estimating equation with working independent covariance structure, BR-GEE:Exch, generalized estimating equation with working exchangeable covariance structure

4 logistic regression models for each possible binary dichotomization (1 vs. 2,3,4,5; 1,2 vs. 3,4,5; 1,2,3 vs. 4,5; 1,2,3,4 vs. 5) of the outcome. Overall, the estimate was smaller for the BR-GEE than for the GEE. And the results from any methods suggested that subjects with the active treatment were likely to have a more favorable response when compared to subjects with the control treatment.

## Discussion

In this paper, we have investigated a BR-GEE to obtain GEE estimates of the proportional odds model for longitudinal ordinal data. The BR-GEE was derived by applying adjusted score equations for the proportional odds model in Kosmidis [13] to the GEE as if it were equivalent to the likelihood score equation. A similar approach for

small-sample bias reduction in the GEE estimate for clustered binary data was investigated in recent studies (Paul and Zhang [5], Mondol and Rahman [6], Geroldinger [7], Gosh et al. [8]).

In the simulations that are reported here, the BR-GEE performs better than the standard GEE for small-sample by providing finite estimates in the presence of quasi-complete separation. This finding is consistent with that of Mondol and Rahman [6].

With the above consideration, in the estimation of the regression parameter when the proportional odds model is applied to longitudinal ordinal data in a small-sample size of 50 or less, it is advantageous to apply the BR-GEE instead of the standard GEE. However, these findings are subject to the following limitations. First, our simulation

setting is limited. In fact, we did not consider simulations with extreme parameter settings where datasets with no observations for a specific category throughout all time points are generated. Such settings would lead to conditional performance evaluation, and therefore we decided to avoid such configurations in this study. Second, we assumed we had complete data, so further study may be required to evaluate the properties of the BR-GEE when there are missing data. In addition, the bias-corrected covariance estimator for GEE with longitudinal ordinal data may also be in need of further investigation. The bias correction methods for the variance-covariance matrix in small-sample GEE estimation have primarily been evaluated assuming binary or count response data. Considering ordinal response data, we believe that new insights can be obtained by applying the previously proposed methods as well as the method proposed in this study and conducting a detailed evaluation of operational characteristics. As alternative models when proportional odds model is not appropriate, we have other types of logit including the adjacent-categories or continuation-ratio. Furthermore, for applying the stereotype model to longitudinal data, estimation methods using GEE was considered [25]. It is worth considering the application of bias-reduction methods to these models.

#### Abbreviations

|        |                                               |
|--------|-----------------------------------------------|
| GEE    | Generalized estimating equations              |
| BR-GEE | Bias-reduced generalized estimating equations |
| MSE    | Mean squared error                            |
| RMSE   | Root mean squared error                       |

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12874-024-02259-6>.

Supplementary Material 1.

#### Acknowledgements

Not applicable.

#### Authors' contributions

Y.T. and T.S. contributed to the study conception and design; Y.T. conducted the simulation study and the statistical analysis; Y.T. wrote the draft of the manuscript; and T.S. provided input. All authors have read and approved the final manuscript.

#### Funding

This research was supported by AMED under Grant/Award Number JP211k0201702. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

#### Availability of data and materials

SAS code for the BR-GEE can be found at <https://github.com/yukiotada0/bргеegor.git>. Data on postoperative pain after laparoscopic cholecystectomy can be found in Lumley [14].

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare no competing interests.

Received: 6 July 2023 Accepted: 7 June 2024

Published online: 28 June 2024

#### References

1. Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis. 2nd ed. New Jersey: John Wiley & Sons; 2011.
2. International Conference on Harmonization. ICH E9(R1) - Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2019. [https://database.ich.org/sites/default/files/E9-R1\\_Step4\\_Guideline\\_2019\\_1203.pdf](https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf). Accessed 25 Mar 2023.
3. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73(1):13–22. <https://doi.org/10.1093/biomet/73.1.13>.
4. Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80(1):27–38. <https://doi.org/10.1093/biomet/80.1.27>.
5. Paul S, Zhang X. Small sample GEE estimation of regression parameters for longitudinal data. *Statist Med*. 2014;33(22):3869–81. <https://doi.org/10.1002/sim.6198>.
6. Mondol MH, Rahman MS. Bias-reduced and separation-proof GEE with small or sparse longitudinal binary data. *Statist Med*. 2019;38(14):2544–60. <https://doi.org/10.1002/sim.8126>.
7. Geroldinger A, Blagus R, Ogden H, et al. An investigation of penalization and data augmentation to improve convergence of generalized estimating equations for clustered binary outcomes. *BMC Med Res Methodol*. 2022;22(168). <https://doi.org/10.1186/s12874-022-01641-6>.
8. Gosh M, Ishii R, Noma H, Maruo K. A comparison of bias-adjusted generalized estimating equations for sparse binary data in small-sample longitudinal studies. *Statist Med*. 2023;42(15):2711–27. <https://doi.org/10.1002/sim.9744>.
9. Bie R, Haneuse S, Huey N, Schildcrout J, McGee G. Fitting marginal models in small samples: a simulation study of marginalized multilevel models and generalized estimating equations. *Statist Med*. 2021;40:5298–312. <https://doi.org/10.1002/sim.9126>.
10. Greenland S. Small-sample bias and corrections for conditional maximum-likelihood odds-ratio estimators. *Biostatistics*. 2000;1(1):113–22. <https://doi.org/10.1093/biostatistics/1.1.113>.
11. Agresti A. Analysis of Ordinal Categorical Data. 2nd ed. New Jersey: John Wiley & Sons; 2010.
12. Kosmidis I, Firth D. Bias reduction in exponential family nonlinear models. *Biometrika*. 2009;96(4):793–804. <https://doi.org/10.1093/biomet/asp055>.
13. Kosmidis I. Improved estimation in cumulative link models. *J Royal Stat Soc Ser B Methodol*. 2014;76(1):169–96. <https://doi.org/10.1111/rssb.12025>.
14. Lumley T. Generalized estimating equations for ordinal data: a note on working correlation structures. *Biometrics*. 1996;52(1):354–61. <https://doi.org/10.2307/2533173>.
15. Heagerty PJ, Zeger SL. ginal regression models for clustered ordinal measurements. *J Am Stat Assoc*. 1996;91(435):1024–36. <https://doi.org/10.1080/01621459.1996.10476973>.
16. Fahrmeir L, Pritscher L. Regression analysis of forest damage by marginal models for correlated ordinal responses. *Environ Ecol Stat*. 1996;3(3):257–68. <https://doi.org/10.1007/BF00453014>.
17. Lipsitz SR, Kim K, Zhao L. Analysis of repeated categorical data using generalized estimating equations. *Statist Med*. 1994;13(11):1149–63. <https://doi.org/10.1002/sim.4780131106>.
18. Liang KY, Zeger SL, Qaqish B. Multivariate regression analyses for categorical data. *J Royal Stat Soc Ser B Methodol*. 1992;54(1):3–40. <https://doi.org/10.1111/j.2517-6161.1992.tb01862.x>.

19. Williamson JM, Kim K, Lipsitz SR. Analyzing bivariate ordinal data using a global odds ratio. *J Am Stat Assoc.* 1995;90(432):1432–7. <https://doi.org/10.2307/2291535>.
20. Hines RJO. Analysis of clustered polytomous data using generalized estimating equations and working covariance structures. *Biometrics.* 1997;53(4):1552–6. <https://doi.org/10.2307/2533523>.
21. Mancl LA, DeRouen TA. A covariance estimator for GEE with improved small-sample properties. *Biometrics.* 2001;57(1):126–34. <https://doi.org/10.1111/j.0006-341x.2001.00126.x>.
22. Wang M, Kong L, Li Z, Zhang L. Covariance estimators for generalized estimating equations (GEE) in longitudinal analysis with small samples. *Statist Med.* 2016;35:1706–21. <https://doi.org/10.1002/sim.6817>.
23. Ibrahim NA, Suliadi S. Generating correlated discrete ordinal data using R and SAS IML. *Comp Method Prog Biomed.* 2011;104(3):e122–32. <https://doi.org/10.1016/j.cmpb.2011.06.003>.
24. Jorgensen JO, Gillies RB, Hunt DR, Caplehorn JR, Lumley T. A simple and effective way to reduce postoperative pain after laparoscopic cholecystectomy. *Aust N Z J Surg.* 1995;65(7):466–9. <https://doi.org/10.1111/j.1445-2197.1995.tb01787.x>.
25. Spiess M, Fernández D, Nguyen T, Liu I. Generalized estimating equations to estimate the ordered stereotype logit model for panel data. *Statist Med.* 2020;39:1919–40. <https://doi.org/10.1002/sim.8520>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.