STATISTICS IN MEDICINE
Statist. Med. 2009; **28**:2338–2355
Published online 26 May 2009 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3622

A comparison of several approaches for choosing between working correlation structures in generalized estimating equation analysis of longitudinal binary data

Justine Shults[1,*,†], Wenguang Sun[1], Xin Tu[2], Hanjoo Kim[1], Jay Amsterdam[3], Joseph M. Hilbe[4,5] and Thomas Ten-Have[1]

[1]*Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, PA 19034, U.S.A.*
[2]*Department of Biostatistics and Computational Biology and Department of Psychiatry, University of Rochester, Rochester, NY 14642, U.S.A.*
[3]*Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA 19034, U.S.A.*
[4]*School of Social and Family Dynamics, Arizona State University, Tempe, AZ 85287, U.S.A.*
[5]*Department of Nutrition, Arizona State University Polytechnic, Mesa, AZ 85212, U.S.A.*

SUMMARY

The method of generalized estimating equations (GEE) models the association between the repeated observations on a subject with a patterned correlation matrix. Correct specification of the underlying structure is a potentially beneficial goal, in terms of improving efficiency and enhancing scientific understanding. We consider two sets of criteria that have previously been suggested, respectively, for selecting an appropriate working correlation structure, and for ruling out a particular structure(s), in the GEE analysis of longitudinal studies with binary outcomes. The first selection criterion chooses the structure for which the model-based and the sandwich-based estimator of the covariance matrix of the regression parameter estimator are closest, while the second selection criterion chooses the structure that minimizes the weighted error sum of squares. The rule out criterion deselects structures for which the estimated correlation parameter violates standard constraints for binary data that depend on the marginal means. In addition, we remove structures from consideration if their estimated parameter values yield an estimated correlation structure that is not positive definite. We investigate the performance of the two sets of criteria using both simulated and real data, in the context of a longitudinal trial that compares two treatments for major depressive episode. Practical recommendations are also given on using these criteria to aid in the efficient selection of a working correlation structure in GEE analysis of longitudinal binary data.

KEY WORDS: generalized estimating equations; longitudinal data; first-order autoregressive correlation structure; correlated binary data; longitudinal study

*Correspondence to: Justine Shults, Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, PA 19034, U.S.A.
†E-mail: jshults@mail.med.upenn.edu

Contract/grant sponsor: NIH-R01CA; contract/grant number: 096885

Received 21 November 2007
Accepted 8 April 2009

1. INTRODUCTION

The method of generalized estimating equations (GEE, [1, 2]) is an extremely popular approach. GEE extends generalized linear models to correlated data by assuming a generalized linear model for the outcome variable and a structured correlation matrix to describe the pattern of association amongst the repeated measurements on each subject, or cluster.

When conducting a GEE analysis, the correlations are often viewed as nuisance parameters. However, it can be beneficial to carefully model the correlation parameters, in order to: (1) avoid a potential substantive loss in efficiency in the estimation of the regression parameter that can result from applying the incorrect working structure, in particular, for larger values of the correlation, small sample sizes [3], and time-varying covariates [4]; (2) avoid problems with respect to infeasibility in estimation of the correlation parameters that can also result from misidentification of the true structure [5, 6]; and (3) enhance scientific understanding, e.g. when the correlations are of substantive interest and a particular working structure is biologically plausible.

In this paper we compare several simple criteria that can be easily obtained in a GEE analysis and used to guide the final selection of a working correlation structure. We consider correlated binary data and assume that the true correlation structure is first-order autoregressive (AR1), where the correlation between measurements j and k on subject i (y_{ij} and y_{ik}, respectively) is given by $\text{Corr}(y_{ij}, y_{ik}) = \alpha^{|j-k|}$. The AR1 structure is often plausible for longitudinal trials with measurements that are approximately equally spaced in time because it forces the correlation between consecutive measurements on a subject to decrease with increasing separation in *measurement occasion*.

In addition to the AR1 structure, we also planned to apply other structures to assess the sensitivity of results to the choice of working structure. These included the exchangeable, for which $\text{Corr}(y_{ij}, y_{ik}) = \alpha$ for $j \neq k$; the tri-diagonal structure, for which $\text{Corr}(y_{ij}, y_{ik}) = \alpha$ for $|j-k| = 1$ or 0 for $|j-k| > 1$; and the identity, for which $\text{Corr}(y_{ij}, y_{ik}) = 0$ for $j \neq k$. The exchangeable model is useful for clustered observations, e.g. students in the same classroom or members of the same household. It might also be plausible in longitudinal studies of short duration, when little decay is expected in the correlation of measurements with increasing separation in time. As noted by a reviewer, in contrast to the exchangeable structure, for which the correlation remains uniformly high across the pairs of time points, only adjacent observations are correlated for the tri-diagonal structure. The tri-diagonal structure could therefore be relevant in longitudinal studies, to investigate properties at this limiting case of minimal correlation among the observations. We considered the tri-diagonal, exchangeable, and AR1 structures because these are the standard working structures that are available in most software packages that implement GEE, e.g. in PROC GENMOD in SAS. We also note that the unstructured correlation matrix could be applied, but this is straightforward only when we have balanced data, i.e. an equal number of observations per subject and subjects per cluster. In addition, an $n \times n$ unstructured correlation matrix involves $n(n-1)/2$ parameters and is therefore not parsimonious. For example, for the longitudinal trial we consider in this paper, some subjects have eight measurements and would therefore require estimation of $8(7)/2 = 28$ parameters, for an unstructured matrix.

We describe and compare several criteria for identification of a working correlation structure in a GEE analysis of longitudinal binary data. We propose *selection criteria* that may be similar for several candidate structures, in which case we also propose *rule out procedures* that may remove

DOI: 10.1002/sim

additional structures from consideration. The approaches we consider for selection of a working correlation structure include criteria by Rotnitzky and Jewell [7] whose implementation was first suggested by Wang and Carey in [8]. The Rotnitzky–Jewell (RJ) criteria were also implemented by Shults *et al*. in [9] in the assessment of the fit of a banded Toeplitz matrix in a GEE analysis of a study of interstitial cystitis in women. We also evaluate a simple selection criterion (SC) proposed by Shults and Chaganty in [10], which was to choose the correlation structure that corresponds to the minimum value of an objective function, the *weighted error sum of squares*. In addition, we propose and evaluate a rule out criterion that is based on Prentice's [11] observation that the correlation parameter must satisfy additional constraints for correlated binary outcomes. The basic idea behind this criterion is to rule out structures that correspond to a violation of bounds for the correlation for binary data that are described in [11]. In addition, we propose a rule out criterion based on Crowder's observation [5] that there may be a problem with respect to the infeasibility and existence of estimates of the correlation parameter when the working structure is misspecified. This final criterion is to rule out structures that correspond to an estimated correlation matrix that is not positive definite.

We evaluate the approaches for the selection of a working correlation structure in the context of a randomized clinical trial that compared two medications for the treatment of major depressive episode (MDE): the standard medication (Lithium) versus the newer treatment (Venlafaxine). Preliminary analysis indicated that the values of MDE on a subject tended to be more similar if they were collected more closely together in time. Because the measurements were approximately equally spaced, we therefore identified the AR1 structure as a biologically plausible correlation structure for this study.

We demonstrate that the criteria considered in this paper were unanimous in their selection of the AR1 structure in the GEE analysis of the binary outcomes from the Venlafaxine trial. In addition, we show (via simulations) that, in particular, the performance of the RJ criteria to correctly identify the true correlation structure for larger values of the correlation, appeared to be superior to the other criteria that we evaluated.

Our outline for this paper is as follows. In Section 2 we define notation, provide the model for analysis of the Venlafaxine study, and briefly review the GEE estimation approach. Next, in Section 3 we describe the approaches we consider for choosing a working correlation structure. Evaluation based on simulations is then provided in Section 4: Section 4.1 describes our approach for simulating data with an AR1 structure; Section 4.2 describes our method for comparison of the selection criteria; Section 4.3 describes the simulation results; and Section 4.4 presents some asymptotic results that explain some of the simulation findings. The different approaches for identification of a working structure are then applied in the analysis of the Venlafaxine study in Section 5. Finally, Section 6 presents a discussion of our findings and our recommendations with regard to selection of a working correlation structure in a GEE analysis of correlated binary data from a longitudinal trial.

2. THE VENLAFAXINE STUDY

In this section we define some notation, provide our model for analysis of the Venlafaxine study, and briefly review GEE.

DOI: 10.1002/sim

2.1. Notation

We assume the usual setup for longitudinal analysis with GEE [1] in which measurements y_{ij} and associated covariates $x_{ij}=(x_{ij1},\ldots,x_{ijp})'$ are collected on subject i at time t_{ij}, for $j=1,\ldots,n_i$ and $i=1,\ldots,m$. The expected value and variance of measurement y_{ij} on subject i can be expressed using a generalized linear model: $E(y_{ij})=g^{-1}(x'_{ij}\beta)=u_{ij}$ and $\mathrm{Var}(y_{ij})=\phi h(u_{ij})$, respectively, where $g^{-1}(\circ)$ is known as the link function; $h(\circ)$ is the variance function; and ϕ is a known or unknown scale parameter. As in [1], we assume that observations on different subjects are independent. Within subjects, the measurements are correlated, with a pattern of association described by the working correlation structure for observations on subject i, $\mathrm{Corr}(Y_i)=R_i(\alpha)$, that depends on correlation parameter α. The covariance matrix of Y_i is then given by $\mathrm{Cov}(Y_i)=\phi A_i{}^{(1/2)}R_i(\alpha)A_i{}^{(1/2)}$, where $A_i=\mathrm{diag}(h(u_{i1}),\ldots,h(u_{in_i}))$.

2.2. Model for analysis of the Venlafaxine study

This study was a prospective, randomized, open-label comparison of venlafaxine versus lithium monotherapy of BP II MDE. The study was conducted using the Principles of Good Clinical Practice Guidelines, with oversight monitoring by the local Office of Human Research at the University of Pennsylvania School of Medicine and by an independent data and safety monitoring board. The main study hypothesis was that venlafaxine monotherapy would be superior to lithium monotherapy with a similar hypomanic switch rate. In this paper we focus on a secondary comparison of change in the probability of MDE over time between the two treatment groups. For more details of the study, see [12].

The outcomes for analysis of the Venlafaxine study were binary MDE scores y_{ij} that took value 1 if subject i had MDE at measurement occasion j and took value 0 otherwise. To compare treatments, we fit a logistic model for the expectation of MDE scores, with $E(y_{ij})=g^{-1}(\delta_{ij})=P_{ij}$, $\delta_{ij}=x'_{ij}\beta$, and

$$\delta_{ij}=\beta_0+\beta_1 I(\text{Venlafaxine})+\beta_2\text{time}+\beta_3 I(\text{Venlafaxine})\times\text{time} \tag{1}$$

where $g^{-1}(\gamma)=\exp(\gamma)/(1+\exp(\gamma))$. In addition, the variance function $h(\mu)=\mu(1-\mu)$ and $\phi=1$. The covariate vector can therefore be expressed as $x'_{ij}=(x_{ij1}, x_{ij2}, x_{ij3}, x_{ij4})$, where $x_{ij1}=1$; $x_{ij2}=I(\text{Venlafaxine})$, which equals 1 for subjects treated with Venlafaxine and 0 for subjects treated with Lithium; $x_{ij3}=\text{time}$, which takes value $1,2,\ldots,n_i$ for subject i; and $x_{ij4}=I(\text{Venlafaxine})\times\text{time}$, a treatment by time interaction term. There were 26 subjects per treatment group, so that $i=1,2,\ldots,52$. The number of measurements per subject n_i ranged from 2 to 8, with a mean of 6.38. The goal of our analysis was to compare the change over time in the probability of MDE between treatment groups; this was accomplished by testing whether the regression parameter β_3 for the interaction term differed significantly from zero.

As described in Section 1, to model the pattern of intra-subject correlations, we identified the AR1 as a biologically plausible correlation matrix. However, we also planned to implement other structures, including the tri-diagonal, identity, and exchangeable, to assess the sensitivity of results to the choice of working structure.

DOI: 10.1002/sim

2.3. Brief review of GEE

GEE is an iterative approach for estimation [1, 2] that alternates between (i) updating the estimate of the regression parameter β by solving the GEE estimating equation for β and (ii) updating the estimate of the correlation parameter α. The GEE estimating equation for β (equation (6) in [1]) depends on the estimate of the working correlation structure for observations on subject i, $\mathrm{Corr}(Y_i) = R_i(\alpha)$. Typically, moment estimates are used for estimation of α. We will use $\hat{\alpha}$ as a generic term whose value will change according to the choice of working correlation structure. To implement GEE in our simulations, we used Stata [13], which applies moment estimates [14] that simplify as follows for $n_i = n$. For the exchangeable structure,

$$\hat{\alpha} = \frac{\sum_{i=1}^{m} \sum_{k \neq j} z_{ik} z_{ij}}{\sum_{i=1}^{m} (n-1) \sum_{j=1}^{n} z_{ij}^2} \tag{2}$$

which is identical to the estimate provided in [6]. For the tri-diagonal structure,

$$\hat{\alpha} = \frac{\sum_{i=1}^{m} n \sum_{k=1}^{n-1} z_{ik} z_{ik+1}}{\sum_{i=1}^{m} (n-1) \sum_{j=1}^{n} z_{ij}^2} \tag{3}$$

For the AR1 structure, Stata implements an algorithm by Newton [15].

The distribution of the GEE estimate of β, $\hat{\beta}$ is asymptotically normal. As discussed in [4], misspecification of the true correlation structure will not typically impact the consistency of $\hat{\beta}$, but may result in efficiency loss in estimation of β. For an excellent text on GEE, see [16].

3. CRITERIA FOR CHOOSING A WORKING CORRELATION STRUCTURE

Here we describe the approaches we compare for choosing between several correlation structures in a GEE analysis of longitudinal binary data; these include *selection criteria* for identification of a structure and supplementary *rule out* criteria that may be helpful in removing structures from consideration.

3.1. Selection criteria

3.1.1. RJ criteria. The motivation of the RJ criteria is that if the working correlation structure is close to the true structure, the model-based estimate $\hat{\Sigma}_m$ of the covariance matrix of $\hat{\beta}$ ((35) in [17], that assumes correct specification) and the 'sandwich' estimate $\hat{\Sigma}_s$ ((32) in [17], that is typically considered to be robust to misidentification) should be similar, so that $Q = \hat{\Sigma}_m^{-1} \hat{\Sigma}_s$ should be close to an identity matrix. In this case the quantities $\mathrm{RJ1} = \mathrm{trace}(Q)/p$ and $\mathrm{RJ2} = \mathrm{trace}(Q^2)/p$, where $Q^2 = Q \times Q$ is obtained by matrix multiplication and p is the dimension of Q, should be close to one in value. In addition, $\mathrm{DBAR} = \sum_j (e_j - 1)^2 = \mathrm{RJ2} - 2\mathrm{RJ1} + 1$, where the e_j are the eigenvalues of Q, should be close to zero because if Q is close to an identity matrix, its eigenvalues should be close in value to 1. Although we anticipated that the three criteria should yield similar results, we evaluated RJ1, RJ2, and DBAR separately. For RJ1 we chose the structure that corresponded to the minimum value of $|\mathrm{RJ1} - 1|$; the same rule was applied to RJ2. For DBAR we chose the structure that corresponded to the minimum absolute value of DBAR.

DOI: 10.1002/sim

3.1.2. SC criterion. Shults and Chaganty [10] proposed the following criterion that chooses the structure that minimizes an objective function: Choose the structure that corresponds to the minimum value of the *weighted error sum of squares* $\sum_{i=1}^{m}(\mathbf{Y_i}-\widehat{\mathbf{E}(\mathbf{Y_i})})'\widehat{\mathrm{Cov}(\mathbf{Y_i})}^{-1}(\mathbf{Y_i}-\widehat{\mathbf{E}(\mathbf{Y_i})})=\sum_{i=1}^{m}\widehat{\mathbf{Z}}_\mathbf{i}'\mathbf{R}_i^{-1}(\hat{\alpha})\widehat{\mathbf{Z}}_i$, where $\mathbf{Z}_i$ is an $n_i\times 1$ vector with jth element equal to $(y_{ij}-P_{ij})/\sqrt{P_{ij}(1-P_{ij})}$ for the logistic model (1) that we consider in this paper. The SC criterion is motivated by the hypothesis that the correct working structure in longitudinal analyses should minimize the error sum of squares weighted by the inverse of the working covariance structure.

3.2. Rule out criteria

3.2.1. Failure of final structure to be positive definite or of GEE to converge. In all longitudinal analyses the estimated correlation parameter $\hat{\alpha}$ must satisfy certain constraints in order for the estimated correlation matrix to be positive definite. For example, the structured correlation matrices we consider in this paper will be positive definite only if α takes value in the following intervals (see [18] for the tri-diagonal structure): (i) $(-1,1)$ for AR1, for which $\mathrm{Corr}(y_{ij},y_{ik})=\alpha^{|j-k|}$; (ii) $(-1/(n_m-1),1)$ for the exchangeable, for which $\mathrm{Corr}(y_{ij},y_{ik})=\alpha$ for $j\neq k$; and (iii) $(-1/c_m,1/c_m)$, where $c_m=2\sin(\pi[n_m-1]/2[n_m+1])$ for the tri-diagonal structure, for which $\mathrm{Corr}(y_{ij},y_{ik})=\alpha$ for $|j-k|=1$ and $n_m=\max_i\{n_i\}$ for $n_i=$ the number of observations on subject i.

Crowder in [5] noted that when the working correlation structure is misspecified, the limiting value of $\hat{\alpha}$ may fail to yield a positive definite correlation matrix. For example, as shown in [6], when the true exchangeable structure is misspecified as tri-diagonal and $n_i=n\forall i$, then the moment estimator for α is still consistent. However, as described above, α must take value in the interval $(-1/c_m,1/c_m)$ in order for the tri-diagonal structure to be positive definite; this interval contains $(-0.5,0.5)$, for all $n_m>2$ and converges to $(-0.5,0.5)$ as $n_m\to\infty$. As a result, if the true value of $\alpha>0.5$ and $n>2$, the limiting value of the GEE moment estimator of α will yield a non-positive definite correlation matrix, when the true exchangeable structure is misspecified as tri-diagonal. It is important to note that this result is not restricted to binary outcomes.

Shults *et al.* [19] consider a GEE analysis in application of the tri-diagonal structure in Stata resulted in the warning that the estimated correlation matrix $R(\hat{\alpha})$ was not positive definite and that GEE failed to converge. In practice, a natural reaction to failure to converge for GEE would be to rule out the structure under consideration because reliable estimates are not available for this structure. In addition, application of the selection criterion considered in this paper is not possible if estimates are not available. We therefore always ruled out a particular working structure when $R(\hat{\alpha})$ was not positive definite or when GEE failed to converge for a particular structure.

In addition, as noted by a reviewer, Dahmen and Ziegler [20] pointed out that GEE will fail to converge if all binary outcomes within a group are identical. The probability of identical outcomes is zero for continuous outcomes; however, for binary outcomes this probability is non-zero and can be large, especially when the number of measurements per subject is small and the correlation is large.

3.2.2. Violation of bounds for correlated binary data. In addition to the constraints required for $R(\hat{\alpha})$ to be positive definite, there are additional restrictions on $\hat{\alpha}$ that must be satisfied when the longitudinal outcomes are binary.

First, GEE provides estimates of the expected values $E(Y_{ij})=P(Y_{ij}=1)=P_{ij}$, the $Q_{ij}=1-P_{ij}$, and the correlation $\mathrm{Corr}(Y_{ij},Y_{ik})=C_{ijk}$ between measurements Y_{ij} and Y_{ik}. Next, the P_{ij}, Q_{ij},

DOI: 10.1002/sim

and C_{ijk} completely determine the bivariate distribution of Y_{ij} and Y_{ik} because, as noted in [11], the pair-wise probabilities $P(Y_{ij}=y_{ij}, Y_{ik}=y_{ik})=P(y_{ij}, y_{ik})$ can be expressed as

$$P(y_{ij}, y_{ik})=P_{ij}^{y_{ij}}\, Q_{ij}^{1-y_{ij}}\, P_{ik}^{y_{ik}}\, Q_{ik}^{1-y_{ik}}\left[1+C_{ijk}\frac{(y_{ij}-P_{ij})(y_{ik}-P_{ik})}{\sqrt{P_{ij}\,P_{ik}\,Q_{ij}\,Q_{ik}}}\right] \tag{4}$$

Prentice [11] pointed out that the probabilities in (4) will be non-negative, i.e. $P(Y_{ij}=y_{ij}, Y_{ik}=y_{ik})\geqslant 0$, only if the correlations satisfy the following constraints (which we refer to as the Prentice constraints) that depend on the marginal means:

$$\text{Lower}_i(j,k)\leqslant \text{Corr}(Y_{ij}, Y_{ik})\leqslant \text{Upper}_i(j,k) \tag{5}$$

where $\text{Lower}_i(j,k)=\max\{-(w_{ij}w_{ik})^{1/2}, -(w_{ij}w_{ik})^{-1/2}\}$, $\text{Upper}_i(j,k)=\min\{(w_{ij}/w_{ik})^{1/2}, (w_{ij}/w_{ik})^{-1/2}\}$, and $w_{ij}=P_{ij}(1-P_{ij})^{-1}$, for $i=1,2,\ldots,m$, $j=1,2,\ldots,n_i$, and $k=1,2,\ldots,n_i$.

The usefulness of the Prentice constraints to aid in ruling out particular working correlation structures is based on the following observations: First, the boundary values L_W and U_W for α are functions of the P_{ij} and therefore of β. Typically, the GEE estimates of β will be consistent, i.e. $\hat{\beta}\xrightarrow{p}\beta$, even when the working structure is misspecified. As a result, the boundary values will be estimated consistently, even when the working structure is misspecified. However, $\hat{\alpha}$ may fail to be consistent under misidentification, as described in [5]. As a result, the bounds (5) can be violated asymptotically when the working structure is misspecified. For example, Section 4.2.2 of [21] demonstrates that asymptotic violation of bounds can be severe when the AR1 structure is misspecified as exchangeable. That the bounds for α will only be violated asymptotically if the working correlation structure is misspecified suggests that a violation of bounds might be used to rule out a particular candidate working correlation structure. In particular, we suggest the following simple algorithm. *Algorithm to Rule out Working Structures Based on a Violation of Bounds for* $\hat{\alpha}$:

1. For each candidate working correlation structure that yields a positive definite estimated correlation matrix, conduct the GEE analysis and obtain the estimates $\hat{\alpha}$ and $\hat{\beta}$ of the correlation and regression parameters.
2. Obtain the boundary values $\hat{L}_W$ and $\hat{U}_W$ (Appendix A) for $\hat{\alpha}$.
3. Calculate the distance from $\hat{\alpha}$ to the closest boundary value, where the distance equals (i) $\hat{\alpha}-\hat{U}_W$ if $\hat{\alpha}>\hat{U}_W$, (ii)$\hat{L}_W-\hat{\alpha}$ if $\hat{\alpha}<\hat{L}_W$, or (iii) 0 if $\hat{L}_W<\hat{\alpha}<\hat{U}_W$.
4. Rule out the working structure that corresponds to the maximum distance that exceeds zero. (More than one structure may be removed, for multiple structures with large distances.)

A reviewer also noted that [22] offered a different relevant interpretation of the potential for violation of the Prentice constraints; we describe and apply their suggested approach in our analysis of the Venlafaxine Study in Section 5.

4. SIMULATIONS

In this section we investigate the numerical performance of each correlation model criterion using simulations. In Sections 4.1 and 4.2 we describe our approach for simulating binary data and method for comparison of the criteria. Results are then discussed in Sections 4.3 and 4.4.

DOI: 10.1002/sim

4.1. Method for generating binary data

Our approach for simulation of correlated binary data with an AR1 structure requires two assumptions. First, we assumed the following Markovian dependence model for the probability of responses $\mathbf{Y}_i = (y_{i1}, y_{i2}, \ldots, y_{in})$ on subject i that was discussed by Liu and Liang in [23] and by Jung and Ahn in [24]:

$$P(Y_{i1} = y_{i1}, \ldots, Y_{in} = y_{in}) = P(Y_{i1} = y_{i1}) \prod_{j=2}^{n} P(Y_{ij} = y_{ij} | Y_{ij-1} = y_{ij-1}) \tag{6}$$

Next, we assumed that the correlation between any two adjacent observations on a subject is constant, so that $\text{Corr}(Y_{ij}, Y_{ij+1}) = \alpha \forall i$ and j. Using induction, it is relatively straightforward to show that assumption of the Markovian dependence model (6) and of constant correlations, does indeed yield data with an AR1 structure for the responses within each subject, so that $\text{Corr}(Y_{ij}, Y_{ik}) = \alpha^{|j-k|}$. The proof is available in Appendix A of [21].

In addition, a reviewer suggested that we run additional simulations for an exchangeable correlation structure; We simulated data for the exchangeable structure according to the approach in [25].

4.2. Approach for simulation-based comparison of criteria for selection of a correlation structure

Here we describe our method for comparison of approaches for selection of a correlation structure in a GEE analysis of correlated binary data.

We simulated data according to the approach described in Section 4.1, for the estimated model (1) from the Venlafaxine trial, assuming β equals its estimated value for the AR1 structure in Section 5 and $\alpha \in (\hat{L}_W, \hat{U}_W) = (-0.0683, 0.7752)$. Recall that $\hat{\alpha}$ must take value in the interval $(\hat{L}_W, \hat{U}_W)$ in order to satisfy the constraints (5) for the correlations for binary outcomes. We assumed $n = 8$ observations per subject and simulated data for $m = 20,\ 40,\ 80$, and 160 subjects, with equally sized Venlafaxine and Lithium treatment groups. We conducted simulations for several values of β and n; the results were similar for all simulations. The values of α ranged from -0.0682 to 0.75, in approximate 0.05 increments.

Our simulations were conducted in Stata 10.0 [13]; these programs utilized the xtgee procedure in Stata for implementation of GEE, coupled with user-written programs for implementation of the criteria for selection of a correlation structure. The programs for the simulations and a detailed description of the simulations according to Figure 1 in [26], with some additional information requested by a reviewer, are available on request.

To compare the methods for selection of a structure (RJ1, RJ2, DBAR, and SC) we plotted the percentage of times (out of 1000 simulation runs) that the working correlation structure under consideration was selected for each criterion versus the true value of α, for sample sizes $m = 20,\ 40,\ 80$, and 160. Each simulated data set was fitted under all three working correlation patterns. The results under assumption of the AR1, exchangeable, and tri-diagonal working structures are displayed in Figures 1–3, respectively. The data were simulated according to an AR1 structure, therefore, Figure 1 displays results when the working structure is correctly assumed to be AR1, while Figures 2 and 3 display results when the working structure is incorrectly assumed to be exchangeable and tri-diagonal, respectively.

To assess the bounds rule out criterion described in Section 3.2.2, we plotted the percentage of times (out of 1000 simulation runs) that this criterion ruled out each working structure versus α, for

DOI: 10.1002/sim

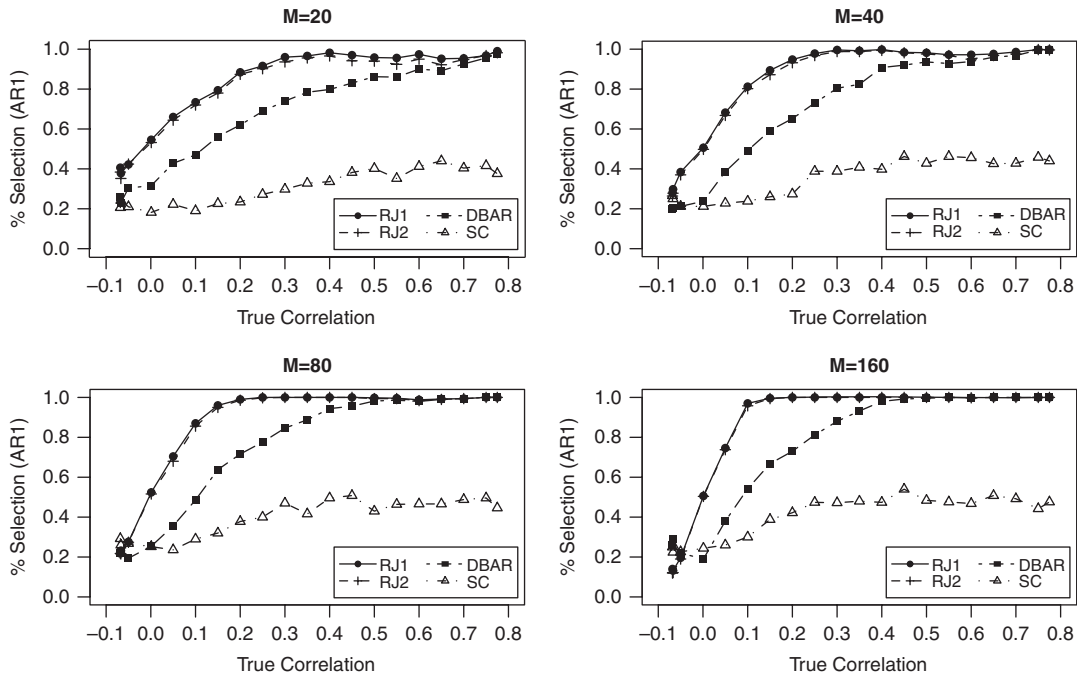

Figure 1. Proportion of 1000 simulation runs in which the AR1 working structure was correctly selected, versus α. Candidate working structures include the AR1; exchangeable, and tri-diagonal. The true structure is AR1. The logistic model for the simulated binary outcomes is (1), with $\beta = (0.6679, 0.4678, -0.2273, -0.282)$. The RJ1 and RJ2 criteria, and to a slightly lesser extent, the DBAR criterion, are superior for $\alpha > 0$. In contrast, the ability of the SC criterion to correctly select the AR1 working structure is relatively weak for all values of α.

$m = 20,\ 40,\ 80$, and 160; the results are displayed in Figure 4. In addition, as noted earlier, because failure of the estimated correlation matrix to be positive definite usually results in non-convergence for GEE, we always ruled out a structure if $R(\hat{\alpha})$ was not positive definite; this criterion was therefore not evaluated separately.

4.3. Results of the simulations

A clear result of the simulations for the true AR1 structure was that the performance of the RJ1 and RJ2 selection criteria (and to a slightly lesser extent, the DBAR criterion) appeared to be superior for $\alpha > 0$. As shown in Figure 1, the RJ1 and RJ2 criteria both did a superior (and virtually indistinguishable) job of correctly selecting the AR1 structure, even for small m and α. For example, even for $m = 20$ subjects, the RJ1 and RJ2 correctly selected the AR1 structure ≈ 100 per cent of the time, for $\alpha > 0.30$. For $m = 40$, the RJ1 and RJ2 approaches correctly selected the AR1 structure ≈ 100 per cent of the time, for $\alpha > 0.20$. The performance of the DBAR approach was also superior, although to a lesser degree than the RJ1 and RJ2 criteria for smaller α.

The performance of the SC criterion was much weaker than that of the Rotnitzky–Jewell (RJ1, RJ2, DBAR) criteria for selection. As shown in Figures 1 and 2, the SC criterion did a poor job of distinguishing between the true AR1 and exchangeable structures, for all values of m and α.

DOI: 10.1002/sim

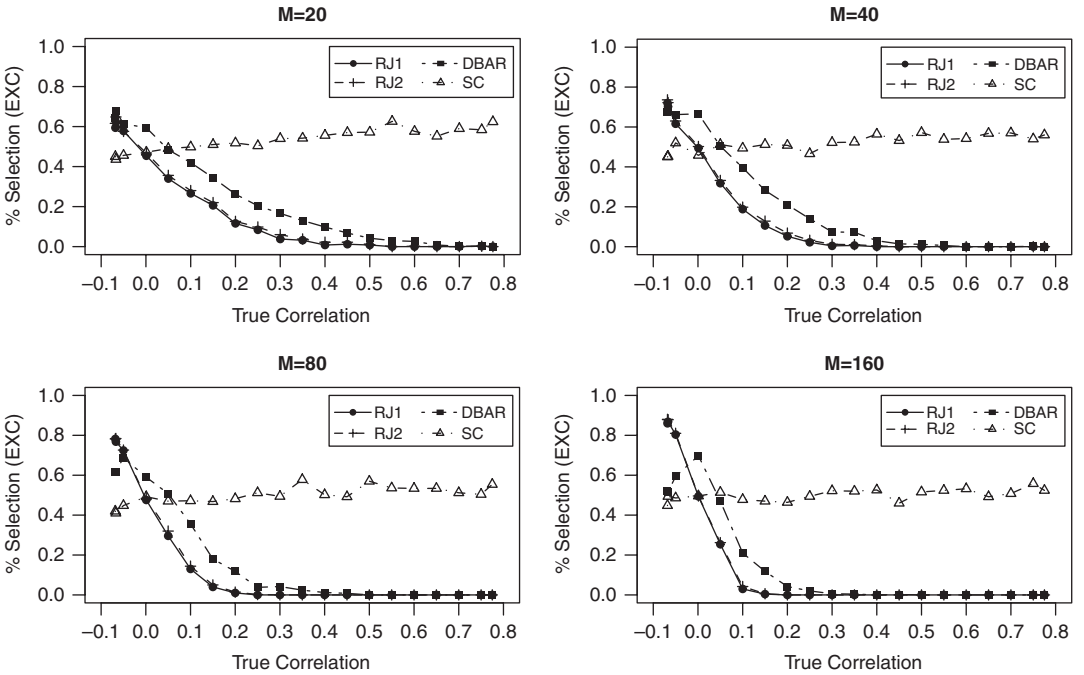

Figure 2. Proportion of 1000 simulation runs in which the exchangeable working structure was incorrectly selected, versus α. Candidate working structures include the AR1; exchangeable, and tri-diagonal. The true structure is AR1. The logistic model for the simulated binary outcomes is (1), with $\beta = (0.6679, 0.4678, -0.2273, -0.282)$. The RJ1, RJ2 and DBAR criteria, and to a slightly lesser extent, the DBAR criterion, perform well for $\alpha > 0$ because they are unlikely to incorrectly select the exchangeable structure. The performance of the SC criterion is much weaker and does not improve with increasing sample size.

However, as shown in Figure 3, all approaches did well with regard to avoiding incorrect selection of the tri-diagonal structure, although the performance of the DBAR and SC criteria was slightly worse than that of RJ1 and RJ2, for smaller α and all sample sizes.

With regard to the bounds rule out procedure, we see (Figure 4) that the bounds approach was very likely to correctly rule out the exchangeable and tri-diagonal structures, for larger values of α and increasing sample sizes. Overall, the bounds approach was very likely to correctly rule out the tri-diagonal structure as α moved from 0; this is indicated by the U-shaped curve for the working tri-diagonal structure in Figure 4. However, the bounds approach was more likely to incorrectly rule out the AR1 structure (in favor of the exchangeable structure) for $\alpha < 0$ and increasing m. In fact, the performance of all methods we considered was weak for negative α. However it is important to note that the lower bound for α was close to zero for our simulation scenario; As a result, the working structures were similar to an identity matrix, and to each other, for $\alpha < 0$.

As noted earlier, a reviewer also requested additional simulations for a true exchangeable correlation structure. Owing to space limitations, the graphs are available on request. The simulation results were similar to those for the true AR1 structure, with two important exceptions. First, as for the true AR1 structure, the performance of RJ1, RJ2, and DBAR was superior to SC with respect to correctly selecting the exchangeable structure and avoiding incorrect selection of the

DOI: 10.1002/sim

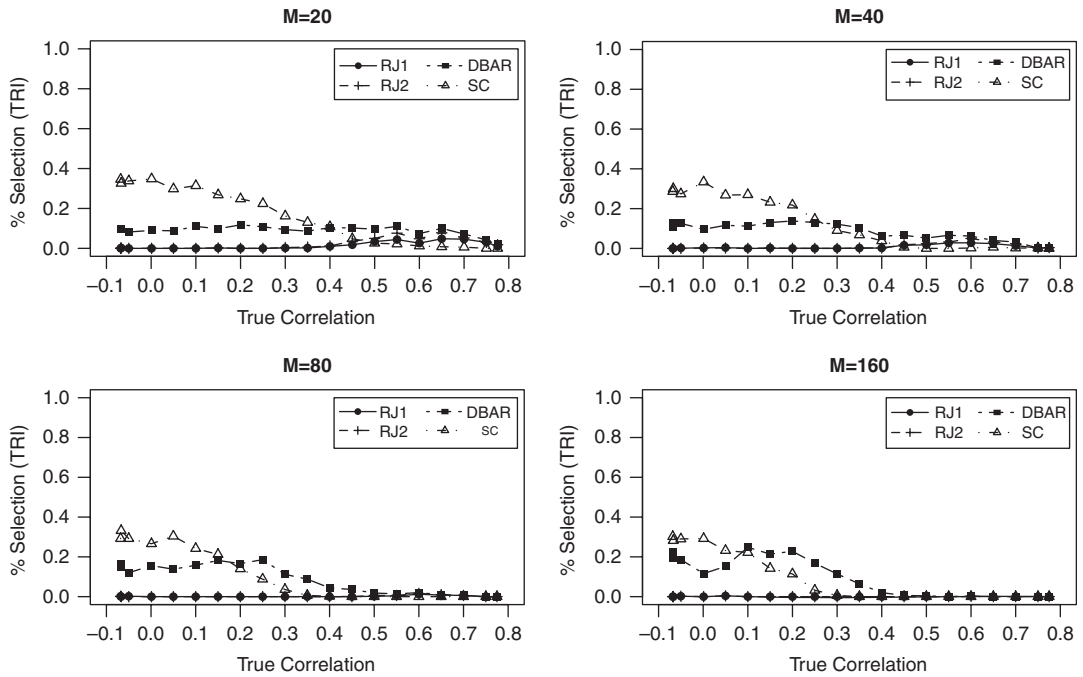

Figure 3. Proportion of 1000 simulation runs in which the tri-diagonal working structure was incorrectly selected, versus α. Candidate working structures include the AR1; exchangeable and tri-diagonal. The true structure is AR1. The logistic model for the simulated binary outcomes is (1), with $\beta = (0.6679, 0.4678, -0.2273, -0.282)$. All methods perform well for this structure because they are unlikely to incorrectly select the tri-diagonal structure. However, the DBAR and SC criteria perform slightly worse than the RJ1 and RJ2 criteria, for smaller α and all sample sizes.

tri-diagonal and AR1 structures; however, the performance of RJ1, RJ2, and DBAR was more similar for the true exchangeable structure, with superior performance for DBAR for negative α. Next, the bounds procedure did not perform well with respect to correctly ruling out the tri-diagonal or AR1 structures in favor of the true exchangeable structure. We will explain the reason the bounds approach failed to correctly rule out the tri-diagonal and AR1 structures in the next section.

4.4. Contrast of asymptotic and numerical results

In this section, we present asymptotic results to provide insight into the simulation results presented in Section 4.3. We considered the model (1) from the Venlafaxine trial and assumed that β equals its estimated value for the AR1 structure presented in Section 5. To compare simulation results for increasing sample sizes, in Table I we display the following asymptotic quantities for each working structure: (i) the limiting values of $\hat{\alpha}$; (ii) the limiting value of the estimated Prentice constraints $(\hat{L}_W, \hat{U}_W)$ that was obtained by calculating (L_W, U_W) for each working structure (see Appendix A) at the assumed value of β for the model (1); (iii) the interval on which α yields a positive definite correlation matrix (see Section 3.2.1); and (iv) the limiting value of $\hat{\alpha}$ in the interval $(L_W, U_W) = (-0.0683, 0.7752)$ for the AR1 structure, on which α takes value for the simulations. The candidate working structures were exchangeable, tri-diagonal, and AR1; see Section 3.2.1 for

DOI: 10.1002/sim

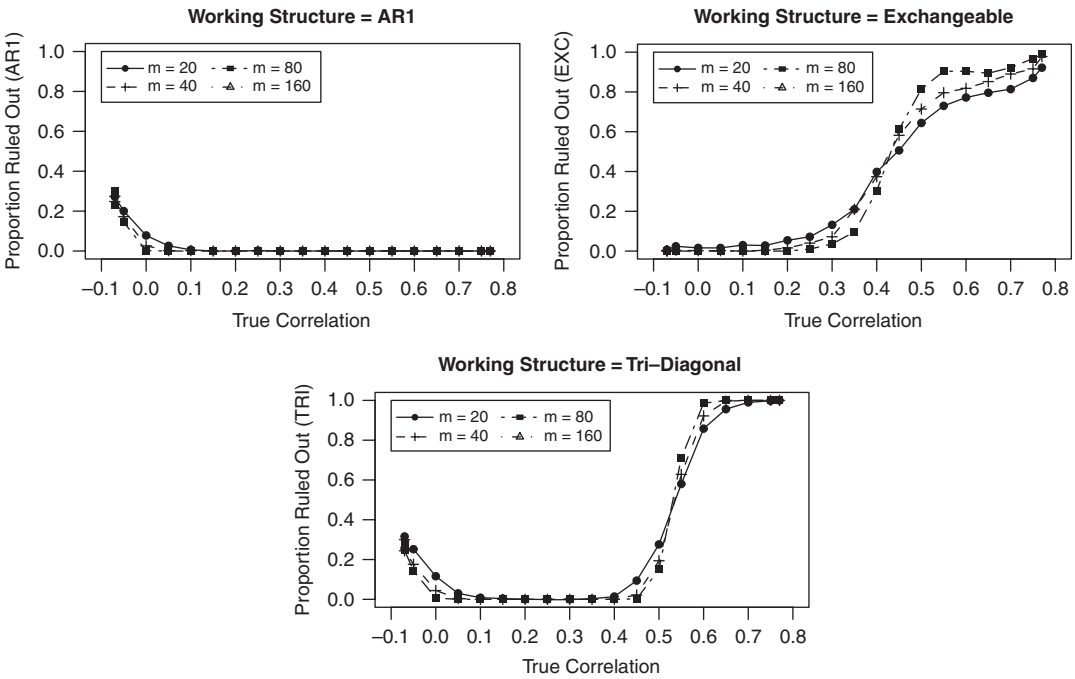

Figure 4. Proportion of 1000 simulation runs in which the bounds approach ruled out each working structure, versus α. Candidate working structures include the AR1; exchangeable, and tri-diagonal. The true structure is AR1. The logistic model is (1), with $\beta=(0.6679, 0.4678, -0.2273, -0.282)$. The bounds approach does a good job of correctly ruling out the tri-diagonal and exchangeable structures for large $\alpha>0$ and increasing m. For small $\alpha<0$ the bounds approach correctly rules out the tri-diagonal structure a high proportion of times; however, it also incorrectly rules out the AR1 structure in favor of the exchangeable structure, for $\alpha<0$.

definitions of these structures. We also note that the identity structure is a special case (for $\alpha=0$) of each of the candidate structures.

As noted by a reviewer, comparing α across correlation structures could be problematic because this parameter could be viewed as having a different interpretation in each structure. To allow for comparison across the models we consider, we note that the off-diagonal elements of the tri-diagonal, AR1, and exchangeable structures are constant, which is equivalent to assuming that the correlation $\text{Corr}(Y_{ij}, Y_{ij+1})$ between adjacent measurements on each subject is constant. In other words, we are selecting from a group of structures that assume the correlation between the jth and $j+1$st measurements on a subject is equal to the correlation between the kth and $k+1$st. We view α as the parameter that represents this constant correlation between adjacent measurements on each subject. Then, for the structures we consider, the correlation between measurements j and k on a subject that are not consecutive is a function $f(\alpha, j, k)$ that depends on the consecutive correlation α and the values of j and k, where $|j-k|>1$; the functions are as follows for each structure: (i) for the AR1 structure, $f(\alpha, j, k)=\alpha^{|j-k|}$; (ii) for the tri-diagonal structure, $f(\alpha, j, k)=0\times\alpha=0$; for the exchangeable structure, $f(\alpha, j, k)=1\times\alpha=\alpha$. Selecting the incorrect structure among the structures that we consider is then equivalent to selecting the incorrect function $f(\alpha, j, k)$, which can have an adverse impact on the estimation of α, as we will see in the results that follow.

DOI: 10.1002/sim

Table I. Limiting values of $\hat{\alpha}$ and of the Prentice constraints $(\hat{L}_w, \hat{U}_w)$ for model (1) when $\beta = (0.6679, 0.4678, -0.2273, 0.282)$; the true correlation structure is AR1; and the working structures are AR1, tri-diagonal (TRI), or exchangeable (EXC).

Working structure	Limiting value of $\hat{\alpha}$	Limiting value of $(\hat{L}_w, \hat{U}_w)^*$	Interval on which α yields a PD matrix	Limiting value of $\hat{\alpha}$ for $\alpha \in$ $(-0.0683, 0.7752)$
AR1	α	$(-0.0683, 0.7752)$	$(-1, 1)$	$(-0.0683, 0.7752)$
TRI	α	$(-0.0683, 0.7752)$	$(-0.531, 0.531)$	$(-0.0683, 0.7752)$
EXC†	$\frac{2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}\alpha^{k-j}}{n(n-1)}$	$(-0.0683, 0.1682)$‡	$(-0.143, 1)$	$(-0.0161, 0.5089)$

In addition, the interval on which α yields a positive definite (PD) correlation matrix is provided for each working structure.

*The limiting value of $(\hat{L}_w, \hat{U}_w)$ is (L_w, U_w), the interval on which α must take value in order to satisfy the bounds for the correlations (5) for binary outcomes.

†The limiting value of $\hat{\alpha}$ when the AR1 structure is misspecified as exchangeable is provided in [6].

‡The limiting value of $\hat{\alpha}$ exceeds the limiting value of $\hat{U}_W$ for $\alpha > 0.4298$.

The values shown in Table I are helpful in explaining some findings of the simulations. For example, if the limiting value of $\hat{\alpha}$ does not take value in the limiting value of $(\hat{L}_W, \hat{U}_W)$ for a particular working structure and interval for α, the bounds approach will be more likely to rule out that structure over the given interval, as the sample size increases. For example, on the basis of Table I, the bounds approach (Figure 4) should be more likely to rule out the exchangeable structure for increasing m and $\alpha > 0.4298$ because the limiting value of $\hat{\alpha}$ exceeds the limiting value of $\hat{U}_W$ for $\alpha > 0.4298$. However, it is also interesting to note that for $\alpha < 0$ the limiting value of $\hat{\alpha}$ exceeds α; e.g. for $\alpha = -0.0683$ the limiting value of $\hat{\alpha}$ is -0.0161 for the exchangeable structure. This fact, coupled with the fact that [21] observed that a small violation of bounds is likely when the working structure is correctly specified as AR1, for smaller m and α very close to the boundary value, explains why we are less likely to observe a violation of bounds (Figure 4) for the exchangeable structure (and hence less likely to rule out the exchangeable structure on the basis of the bounds approach) for α close to the lower boundary value $L_W = -0.0683$.

In addition, if the limiting value of $\hat{\alpha}$ does not yield a positive definite correlation matrix for a particular structure and interval of α, all approaches will be less likely to select that structure over the given interval, as the sample size increases; this follows from the fact that we always ruled out a working correlation structure that resulted in a non-positive definite matrix, because this typically results in non-convergence for GEE. Table I therefore explains why the tri-diagonal structure was ruled out for increasing m and $\alpha > 0.531$, because the limiting value of $\hat{\alpha}$ yields a non-positive-definite correlation matrix for $\alpha > 0.531$.

As noted earlier, the bounds approach was unlikely to correctly rule out the tri-diagonal or the AR1 correlation structures, in additional simulations that were requested by a reviewer for a true exchangeable structure; this was due to the fact that for model (1) and the assumed value of β, the Prentice constraints for α (i.e. the limiting value of the estimated Prentice constraints) were $(-0.03, 0.1676)$ for the exchangeable structure, versus $(-0.0683, 0.7752)$ for the AR1 structure. In addition, as shown in [6], when the exchangeable structure is misspecified as tri-diagonal or AR1, $\hat{\alpha}$ will still be consistent. Therefore, there will not be an asymptotic violation of the Prentice constraints for α, when the exchangeable structure is misspecified as AR1 or tri-diagonal. This

DOI: 10.1002/sim

highlights an important limitation of the bounds approach, which is that for increasing sample sizes, it will tend to correctly rule out working structures only when misspecification results in an asymptotic violation of the Prentice constraints for α.

5. ANALYSIS OF THE VENLAFAXINE STUDY

Table II displays the results of the GEE analysis for model (1) that was conducted using Stata for implementation of GEE and user-written programs in Stata to obtain the RJ and SC criteria and the boundary values for $\hat{\alpha}$. No results are displayed for the tri-diagonal structure because the estimated correlation matrix was not positive definite, which resulted in a failure for GEE to converge for this structure.

First, we note that the RJ1, RJ2, DBAR, and SC criteria all selected the AR1 as the appropriate patterned correlation matrix because the SC criterion was smallest and (RJ1, RJ2, DBAR) were closest to $(1, 1, 0)$ for this structure. Next, we note that the bounds approach ruled out the exchangeable correlation structure because $\hat{\alpha}=0.4078$ and $(\hat{L}_W, \hat{U}_W)=(-0.0484, 0.1681)$, so that $\hat{\alpha} \notin (\hat{L}_W, \hat{U}_W)$ for this structure.

Table II. Analysis of the Venlafaxine Study for three working structures: Exchangeable (EXC); AR1; and Independent (IND).

	EXC	AR1	IND
$\hat{\alpha}$	0.4078	0.5434	0
$(\hat{L}_W, \hat{U}_W)$	$(-0.0484, 0.1681)$	$(-0.0683, 0.7752)$	NA
RJ1	1.7877	1.2031	2.5723
RJ2	4.3246	1.6830	8.2379
DBAR	1.7493	0.2768	4.0933
SC	397.6683	338.9252	349.8191
$\hat{\beta}_0$	0.6880	0.6679	0.6589
Std.Err.$(\hat{\beta}_0)$	0.4226	0.4214	0.4489
p-value	0.103	0.113	0.142
$\hat{\beta}_1$	0.1038	0.4678	0.3233
Std.Err.$(\hat{\beta}_1)$	0.6337	0.5963	0.6042
p-value	0.870	0.433	0.593
$\hat{\beta}_2$	−0.2682	−0.2273	−0.2381
Std.Err.$(\hat{\beta}_2)$	0.0928	0.0899	0.1025
p-value	0.004	0.011	0.020
$\hat{\beta}_3$	−0.2412	−0.2820	−0.2469
Std.Err.$(\hat{\beta}_3)$	0.1867	0.1667	0.1679
p-value	0.196	0.091	0.141

DOI: 10.1002/sim

The analysis results were similar for all the three structures. However, there was one potentially interesting difference. The interaction parameter β_3 was of primary interest in model (1) because if it differed significantly from zero and had a negative estimated value, this would indicate that the change in the probability of MDE over time is significantly lower for subjects treated with Venlafaxine versus Lithium. The parameter β_3 did not differ significantly from 0 at the 0.05 significance level for any structure; however, it did differ significantly from 0 at the 0.10 level for the AR1 structure. In addition, the estimated value of β_3 was negative and greatest (in absolute value) for the AR1 structure. Although the difference was not striking, application of the AR1 structure relative to application of the other structures was more suggestive of a beneficial treatment effect with Venlafaxine. As suggested by a reviewer, we plotted the trends for the groups to see how they differ in their predictions. These graphs, available on request, suggest that the probability of occurrence of MDE decreased more rapidly over time for the Venlafaxine versus Lithium treatment groups. In addition, the probability of MDE was slightly lower at all measurement occasions post baseline for the exchangeable structures, than for the AR1 and identity structures.

As mentioned earlier, a reviewer also noted that [22] offer a different relevant interpretation of the potential for violation of the Prentice bounds for α. These authors noted that the Prentice bounds can be extremely restrictive under an assumption of an exchangeable structure, and can in some cases reduce to a single point (zero). As a result, they proposed guidelines for selection of the best working structure $R(\alpha)$, which they viewed as a *weight matrix* in the GEE estimating equation for β. In Section 7 of [22] they suggest application of an AR1 matrix for longitudinal data, with $\hat{\alpha} \approx 0$ for weakly dependent binary data; $\hat{\alpha} \in (0.2-0.3)$ for moderately dependent binary data; or $\hat{\alpha} \in (0.4-0.7)$ for strongly dependent binary data. Weak, moderate, or strong dependence should be decided on the basis of descriptive analysis; for example, 'the strongest dependence is indicated if frequencies concentrate near the vectors of all 0 s and all 1s'. Alternatively, they suggested that we could 'apply IEEs (independence estimating equations) first, and then check the (Prentice) bounds ... for each pair (of observations) ... and decide on an appropriate value for α in the midrange of the bounds'. We note that application of IEEs involves solving the GEE estimating equation for β when $R(\alpha)$ is an identity matrix, which is equivalent to conducting a logistic regression for our analysis.

To apply the suggested approach of [22] in our analysis, we first note that since our study is longitudinal, their rules require application of an AR1 structure, which we should term as an AR1 *weight* matrix. Next, descriptive analysis suggested that the correlation could be viewed as strong, although this assessment was somewhat difficult due to the relatively small sample size for this study. Alternatively, if we fit IEEs, the Prentice constraints on α are $(-0.0702, 0.1831)$, with mid-point 0.0564. The suggested estimate $\hat{\alpha}$ of α according to [22] would therefore be 0.0564, or $\alpha \in (0.4, 0.7)$, for an AR1 structure.

We note that our estimated value of $\hat{\alpha} = 0.5434$ for the AR1 structure is within the interval $(0.4, 0.7)$, so that the final estimates based on a GEE analysis are compatible with the suggested approach of [22]. However, our approach might be easier to implement because it yields a unique estimate of $\hat{\alpha}$ for the final working structure. In addition, it allows for consideration of more than one potential structure for longitudinal data. Although the AR1 structure is plausible in many biological studies, application of an exchangeable (or other type) of structure could sometimes be reasonable, especially in a study of very short duration. In addition, the rules in [22] permit a maximum value for $\hat{\alpha}$ of 0.70, which is a limitation if the true value of α exceeds 0.70. A comparison of our approach with application of the rules in [22] for a simulated data set for $\alpha = 0.75$ is available on request.

DOI: 10.1002/sim

6. DISCUSSION

The methods for identification of the most appropriate pattern for the correlations that we considered in this paper all confirmed our choice of the AR1 matrix as the final working correlation structure in our analysis of the Venlafaxine study. In particular, the Prentice constraints (5) were violated for the exchangeable structure in this analysis. As noted by a reviewer, the major software packages that implement GEE, e.g. PROC GENMOD in SAS and xtgee in Stata, do not check and provide a warning for a violation of these constraints. It is therefore important to check that they are satisfied. As noted by Rochon in [27], 'the practitioner must be aware of these restrictions, particularly at the design stage'. For example, when doing power calculations for clinical trials with binary outcomes, it is important to do the computations for values of α that satisfy the Prentice constraints.

During the analysis phase of a study, we suggest consideration of the selection and rule out criteria as well as biological plausibility of a candidate working correlation structure. As we did in Section 5, we suggest fitting all methods for selection of a working structure that we considered in this paper. As noted by a reviewer, the standard errors of all the β coefficients were smaller under the AR1 model that was selected as the correct correlation structure. This demonstrates that the analysis results are sensitive to the choice of working correlation structure. Calculation of the criteria is straightforward and programs in Stata are available on request from the authors. Each set of criteria (rule out or selection) has its own strength in providing useful information in favor of or against a particular correlation structure, as demonstrated using both simulated and real data in Sections 4 and 5. If the criteria are not unanimous in identifying a biologically plausible structure as the correct structure (as they were in our analysis of the Venlafaxine study in Section 5), more careful analysis will be required. For example, more weight might be given to the RJ1, RJ2, and DBAR criteria, due to their superior performance in our simulation study for $\alpha>0$.

A reviewer also noted that it is perhaps not surprising that the SC model performed poorly. The residual is a function of $\hat{\beta}$, and we know that $\hat{\beta}$ is robust to the choice of the working correlation pattern, in the sense that $\hat{\beta}$ will typically be estimated consistently, even when the working correlation structure is misspecified. Thus, any discriminating power for $\mathrm{Cov}(\hat{\beta})$ was, in a sense, 'washed out' by the lack of discriminating power in $\hat{\beta}$ and therefore the residual.

Owing to considerations with regard to paper length, in this paper we considered data with an AR1 structure that is often applied in longitudinal trials with measurements that are equally spaced in time. As described earlier, we made additional comparisons for an exchangeable structure that are available on request. Additional comparisons for other true underlying structures will be the focus of future work. For example, as noted by a reviewer, the mixture of exchangeable and AR1 structures is quite often used in econometrics. In addition, comparisons are planned with the approaches of Pan [28] and Pan and Connet [29].

APPENDIX A: BOUNDS ON α FOR THE AR1, TRI-DIAGONAL, AND EXCHANGEABLE STRUCTURES FOR LONGITUDINAL BINARY OUTCOMES

As noted in Section 3.2.1, the correlation parameter α must satisfy certain constraints, in order for the AR1, tri-diagonal, and exchangeable structures to be positive definite. Then, as described in Section 3.2.2, there are additional constraints $L_W \leqslant \alpha \leqslant U_W$ that depend on the choice of working correlation structure, that must be satisfied for correlated binary data. Here we give the values of L_W and of U_W

DOI: 10.1002/sim

for the particular working structures that we consider in this manuscript: (i) for the exchangeable structure, $L_W = \max_{i,j,k}\{\text{Lower}_i(j,k)\}$ and $U_W = \min_{i,j,k}\{\text{Upper}_i(j,k)\}$. (ii) For the AR1 and tri-diagonal structures, $L_W = \max_{i,j,k:|j-k|=1}\{\text{Lower}_i(j,k)\}$ and $U_W = \min_{i,j,k:|j-k|=1}\{\text{Upper}_i(j,k)\}$. Note that the Prentice constraints are identical for the AR1 and tri-diagonal structures because Theorem 1 of [30] establishes that we only need to check that the constraints are satisfied for adjacent marginal means for the AR1 structure. However, as described in Section 3.2.1, asymptotically in n, α must take value in $(-\frac{1}{2}, \frac{1}{2})$ in order for the tri-diagonal structure to be positive definite, versus $(-1, 1)$ for the AR1 structure. As a result, when we consider the intersection of the Prentice constraints in (5) with the constraints required to be positive definite, the bounds on α will typically be tighter for the tri-diagonal versus AR1 structure.

See Section 2.2 of [21] for further details and simplification for the logistic model and the equicorrelated structure. For the AR1 (and tri-diagonal) structures, [30] proves that the Prentice constraints in (5) for the AR1 (and tri-diagonal structures) for the logistic model are given by $(L_{\text{AR1}}, U_{\text{AR1}})$, where $L_{\text{AR1}} = \max_{i,j}\{-\exp(-|(x_{ij} + x_{ij+1})'\beta|)/2\}$; and $U_{\text{AR1}} = \min_{i,j}\{\exp(-|(x_{ij} - x_{ij+1})'\beta)|/2\}$.

ACKNOWLEDGEMENTS

We would like to acknowledge the Stanley Medical Research Institute Grant Number 01-005 'Antidepressant Monotherapy for Biopolar II Major Depression' that provided the data for the Venlafaxine study in this paper. The ClinicalTrials.gov Identifier for this study is *NCT*00641927. We also thank Shannon Chuai for constructing the figures in R. We are also extremely grateful to the two reviewers for their insightful comments and suggestions.

REFERENCES

1. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**:13–22.
2. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; **42**: 121–130.
3. Albert PS, McShane LM. A generalized estimating equations approach for spatially correlated binary data: applications to the analysis of nueroimaging data. *Biometrics* 1995; **51**:627–638.
4. Sutradhar BC, Das K. On the accuracy of efficiency of estimating equation approach. *Biometrics* 2000; **56**: 622–625.
5. Crowder M. On the use of a working correlation matrix in using generalised linear models for repeated measures. *Biometrika* 1995; **82**:407–410.
6. Wang YG, Carey VJ. Working correlation misspecification, estimation and covariate design: implications for generalized estimating equation performance. *Biometrika* 2003; **90**:29–41.
7. Rotnitzky A, Jewell NP. Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika* 1990; **77**:485–497.
8. Wang YG, Carey VJ. Unbiased estimating equations from working correlation models for irregularly timed repeated measures. *Journal of the American Statistical Association* 2004; **99**:845–852.
9. Shults J, Mazurick CA, Landis JR. Analysis of repeated bouts of measurements in the framework of generalized estimating equations. *Statistics in Medicine* 2006; **25**(23):4114–4128.
10. Shults J, Chaganty NR. Analysis of serially correlated data using quasi-least squares. *Biometrics* 1998; **54**: 1622–1630.
11. Prentice RL. Correlated binary regression with covariates specific to each binary observation. *Biometrics* 1988; **44**:1033–1048.
12. Amsterdam J, Shults J. Comparison of short-term venlafaxine versus lithium monotherapy of bipolar II major depressive episode: a randomized open label study. *The Journal of Clinical Psychopharmacology* 2008; **28**(2): 171–181.

DOI: 10.1002/sim

13. StataCorp. *Stata Statistical Software*: *Release 10*, StataCorp LP, College Station, TX, 2007.
14. Ziegler A, Gromping U. The generalised estimating equations: a comparison of procedures available in commercial statistical software packages. *Biometrical Journal* 1998; **40**:245–260.
15. Newton HJ. *TIMESLAB*: *A Time Series Analysis Laboratory*. Brooks/Cole: Belmont, CA, 1988.
16. Hardin JW, Hilbe JM. *Generalized Estimating Equations*. Chapman & Hall/CRC: London, Boca Raton, 2003.
17. Ziegler A, Arminger G. *Analyzing the Employment Status with Panel Data from GSOEP. A Comparison of the MECCOSA and the GEE1 Approach for Marginal Models*, Deutsches Institut for Wirtschaftsforschung, 1995; 72–80.
18. Shults J. The analysis of unbalanced and unequally spaced longitudinal data using quasi-least squares. *Ph.D. Thesis*, Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia, 1996.
19. Shults J, Ratcliffe SJ, Leonard M. Improved generalized estimating equation analysis via xtqls for quasi-least squares in Stata. *The Stata Journal* 2007; **7**(2):147–166.
20. Dahmen G, Ziegler A. Independence estimating equations for controlled clinical trials with small sample sizes. *Methods of Information in Medicine* 2006; **45**:430–434.
21. Shults J, Sun W, Tu X. On the violation of bounds for the correlation in generalized estimating equation analyses of binary data from longitudinal trials. *UPenn Biostatistics Working Papers. Working Paper 8*, 2006. Available from: http://biostats.bepress.com/upennbiostat/papers/art8.
22. Chaganty NR, Joe H. Efficiency of generalized estimating equations for binary responses. *Journal of the Royal Statistical Society*, *B* 2004; **66**:851–860.
23. Liu G, Liang KY. Sample size calculation for studies with correlated observations. *Biometrics* 1997; **53**:937–947.
24. Jung SH, Ahn WW. Sample size for a two-group comparison of repeated binary measurements using GEE. *Statistics in Medicine* 2005; **24**:2583–2596.
25. Qaqish F. A family of multivariate binary distributions for simulating correlated binary variables with specified marginal means and correlations. *Biometrika* 2003; **90**:455–463.
26. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Statistics in Medicine* 2006; **25**:4279–4292.
27. Rochon J. Application of GEE procedures for sample size calculations in repeated measures experiments. *Statistics in Medicine* 1998; **17**:1643–1658.
28. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics* 2001; **57**:120–125.
29. Pan W, Connet J. Selecting the working correlation structure in generalized estimating equations with application to the lung health study. *Statistica Sinica* 2002; **12**:475–490.
30. Kim H, Hilbe JM, Shults J. On the designation of the patterned associations for longitudinal Bernoulli data: weight matrix versus true correlation structure. *UPenn Biostatistics Working Papers. Working Paper 26*, 2008. Available from: http://biostats.bepress.com/upennbiostat/papers/art26.

DOI: 10.1002/sim