University of Pennsylvania
UPenn Biostatistics Working Papers

Year 2009 *Paper* 33

Quasi-Least Squares with Mixed Linear Correlation Structures

Jichun Xie* Justine Shults† Jon Peet‡
Dwight Stambolian** Mary F. Cotch††

*University of Penn, jichun@mail.med.upenn.edu
†Univeristy of Pennsylvania Department of Biostatistics, jshults@mail.med.upenn.edu
‡jonpeet@comcast.net
**University of Penn, stamboli@mail.med.upenn.edu
††National Institutes of Health, mfc@nei.nih.gov

http://biostats.bepress.com/upennbiostat/art33

Quasi-Least Squares with Mixed Linear Correlation Structures

Jichun Xie, Justine Shults, Jon Peet, Dwight Stambolian, and Mary F. Cotch

Abstract

Quasi-least squares (QLS) is a two-stage computational approach for estimation of the correlation parameters in the framework of generalized estimating equations (GEE). We prove two general results for the class of mixed linear correlation structures: namely, that the stage one QLS estimate of the correlation parameter always exists and is feasible (yields a positive definite estimated correlation matrix) for any correlation structure, while the stage two estimator exists and is unique (and therefore consistent) with probability one, for the class of mixed linear correlation structures. Our general results justify the implementation of QLS for particular members of the class of mixed linear correlation structures that are appropriate for the analysis of familial data, with families that vary in size and composition. We describe the familial structures and implement them in an analysis of optical spherical values in the Old Order Amish (OOA). For the OOA analysis, we show that we would suffer a substantial loss in efficiency, if the familial structures were the true structures, but were misspecified as simpler approximate structures. We also provide software for implementation of the familial structures in R. Key-Words: Quasi-least squares; linear correlation structure; mixed correlation structure; familial data.

1. Introduction

We consider a secondary analysis of optical spherical values in a study in the Old Order Amish (OOA) (Wojciechowski et al., 2009). The families in the OOA study varied in both size and composition, because some nuclear families contained only siblings, while other families included siblings and one or both parents. The goal of the OOA analysis was to relate the expected spherical values, measurements that reflect quality of vision, with gender and age, while also adjusting for the correlation among measurements within each family. Because the correlations in the OOA study were thought to vary according to familial relationship, it was important to allow the sibling-sibling, sibling-father, and sibling-mother correlations to vary in value.

To model the pattern of association amongst measurements in families in the OOA study, we implemented slight generalizations of familial correlation structures considered by Karlin, Cameron, and Williams (1981) and Gleseer (1992). Gleseer (1992) noted that it is computationally difficult to obtain maximum likelihood (ML) estimates of the correlation parameters for normal data, when family sizes are not constant. Gleseer (1992) therefore obtained ML estimates that were weighted averages of estimates obtained for sub-groups with families of equal size. One limitation of the approaches of both Karlin et al. (1981) and of Gleseer (1992) was that they assumed that the expected value of the outcome variable was constant between the siblings. However, it is important to note that Karlin et al. (1981) and Gleseer (1992) allowed the variance of the outcome variable to vary between parents and siblings, while we assume a constant standard deviation of spherical values for all subjects.

We implement the familial correlation structures for analysis of the OOA study with quasi-least squares (QLS). QLS is an approach based on GEE that estimates the correlation parameters in two stages. In the following summary, estimates of the correlation parameters are defined to be feasible if they yield positive definite correlation matrices. Chaganty (1997) considered balanced data and established feasibility of the stage one estimates for the first order auto-regressive AR(1), exchangeable, and tri-diagonal structures. Shults (1996) and Shults and Chaganty (1998) proved feasibility for the afore-mentioned structures, in addition to the Markov structure, for unbalanced data. However, although the stage one estimates exist and are feasible, they are not consistent. Chaganty and Shults (1999) therefore introduced a second stage of QLS and established consistency of the stage two estimates for the AR(1), Markov, and tri-diagonal correlation structures. The second stage of QLS updates the stage one estimate of $\boldsymbol{\alpha}$ by obtaining a solution to an estimating equation (stage two estimating equation for $\boldsymbol{\alpha}$) with an estimating function that only depends on $\boldsymbol{\alpha}$ and the stage one estimate of $\boldsymbol{\alpha}$. Theorem (3.2) of Chaganty and Shults (1999) establishes that if there exists a unique solution to the stage two estimating equation for $\boldsymbol{\alpha}$ that is a continuous and one to one function of the stage one estimate,

then that solution will be consistent for $\boldsymbol{\alpha}$. Software for implementation of QLS is available in SAS (Kim and Shults, 2008), Stata (Shults, Ratcliffe, and Leonard, 2007), MATLAB (Ratcliffe and Shults, 2008), and R (Xie and Shults, 2009).

In this manuscript, we prove two general results for QLS that can be used to justify implementation of QLS for the familial structures we consider. First, we prove that the QLS stage one estimate of $\boldsymbol{\alpha}$ will exist and is feasible with probability one, for any correlation structure. Next, for the class of mixed linear correlation structures, we prove the existence and uniqueness of the QLS stage two estimates, both of which are required for consistency of $\hat{\boldsymbol{\alpha}}$. A benefit of our results is that not only do they justify implementation of QLS for the familial structures we consider in this manuscript, but they can also be used to justify QLS for other structures. For example, Shults, Mazurick, and Landis (2006) implemented QLS for a banded Toeplitz (BT) correlation structure, but did not provide proofs regarding the existence and uniqueness of solutions of the QLS estimating equations for $\boldsymbol{\alpha}$ for this structure. The BT structure is a member of the class of linear correlation structures, so that the results provided in this paper establish the consistency of the QLS estimators of $\boldsymbol{\alpha}$ for this structure. In general, our results for stage one are applicable to any correlation structure, while our results for stage two are applicable to any mixed linear correlation structure.

As an outline for our paper, in Section 2, we give some notation; describe the familial structures we consider; and define mixed linear correlation structures. In Section 3 we then extend QLS for mixed linear correlation structures by proving several results for these structures. Next, we demonstrate the benefit of fitting mixed linear correlation structures: In Section 4, we conduct asymptotic relative efficiency (ARE) comparisons to show that the loss in efficiency in estimation of the regression parameter could be substantial in a QLS (or GEE) analysis of the OOA study, if the true mixed linear correlation structures were misspecified as a simpler, approximate structure. In Section 5 we then present our analysis of the OOA study that demonstrates application of the mixed correlation structure with QLS. The proofs of our theorems and lemmas are provided in the appendices.

2. Background

2.1. **Notation.** We assume that outcomes $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{in_i})^{\mathrm{T}}$ and associated covariates $\boldsymbol{X}_{ij} = (X_{ij1}, \ldots, X_{ijp})^{\mathrm{T}}$ are collected on family i, for $i = 1, \ldots, m$. The expected value and variance of measurement Y_{ij} can be expressed using a generalized linear model (GLM):

$$\mathrm{E}(Y_{ij}) = g^{-1}(X_{ij}^{\mathrm{T}}\beta) = \mu_{ij} \text{ and } \mathrm{Var}(Y_{ij}) = \phi h(\mu_{ij}) \tag{1}$$

respectively, where $g^{-1}(\cdot)$ is the link function; $h(\cdot)$ is the variance function; and ϕ is a known or unknown scale parameter. We assume that observations from different families are independent. However, measurements within families are correlated, with a pattern of association that can be described with

correlation structures for each family i, $\text{Cor}(Y_i) = R_i(\boldsymbol{\alpha})$, that depend on correlation parameter $\boldsymbol{\alpha}$. The covariance matrix of $\boldsymbol{Y}_i$ is then given by $\text{Cov}(\boldsymbol{Y}_i) = \phi A_i^{1/2} R_i(\boldsymbol{\alpha}) A_i^{1/2}$, where $A_i = diag(h(\mu_{i1}), \ldots, h(\mu_{in_i}))$.

2.2. **Familial Structures in the Class of Linear and Mixed Correlation Structures.** Define $\boldsymbol{e}_i$ as the unit vector with only the jth entry equal to 1. We refer to a correlation matrix as **linear** if

$$R_i(\boldsymbol{\alpha}) = \sum_{j=1}^{s} (R_i(\boldsymbol{e}_i) - R_i(0))\alpha_j + R_i(0), \tag{2}$$

so that each element of the matrix can be expressed as a linear combination of $\boldsymbol{\alpha}$. In this case $\boldsymbol{\alpha}$ is identifiable if and only if

$$\sum_{j=1}^{s} (R_i(\boldsymbol{e}_i) - R_i(0))c_j = 0 \text{ if and only if } \boldsymbol{c} = (c_1, \ldots, c_s) = 0. \tag{3}$$

Several linear correlation structures were considered for analysis of the OOA study, which included two-generation families that varied in both size and composition. We assumed that the father-mother, father-sibling, mother-sibling, sibling-sibling correlations were γ, ρ_1, ρ_2 and α, respectively. If family i included both parents and siblings, this resulted in an *extended familial correlation structure* R_i to describe the pattern of association among the n_i measurements on family i:

$$\text{Cor}(Y_i) = \begin{pmatrix} 1 & \gamma & \rho_1 & \rho_1 & \ldots & \rho_1 \\ \gamma & 1 & \rho_2 & \rho_2 & \ldots & \rho_2 \\ \rho_1 & \rho_2 & 1 & \alpha & \ldots & \alpha \\ \rho_1 & \rho_2 & \alpha & 1 & \ldots & \alpha \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \rho_2 & \alpha & \alpha & \ldots & 1 \end{pmatrix}_{n_i \times n_i} . \tag{4}$$

For a family with only a father and siblings, R_i would have a *familial structure*:

$$\text{Cor}(Y_i) = \begin{pmatrix} 1 & \rho_1 & \rho_1 & \ldots & \rho_1 \\ \rho_1 & 1 & \alpha & \ldots & \alpha \\ \rho_1 & \alpha & 1 & \ldots & \alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \alpha & \alpha & \ldots & 1 \end{pmatrix}_{n_i \times n_i} . \tag{5}$$

For a family with only a mother and siblings, R_i would still have a *familial structure*, but with ρ_1 replaced by ρ_2 in (5):

$$\text{Cor}(Y_i) = \begin{pmatrix} 1 & \rho_2 & \rho_2 & \dots & \rho_2 \\ \rho_2 & 1 & \alpha & \dots & \alpha \\ \rho_2 & \alpha & 1 & \dots & \alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_2 & \alpha & \alpha & \dots & 1 \end{pmatrix}_{n_i \times n_i} . \tag{6}$$

Finally, for families with only siblings, the correlation structure would be *exchangeable*:

$$\text{Cor}(Y_i) = \begin{pmatrix} 1 & \alpha & \dots & \alpha \\ \alpha & 1 & \dots & \alpha \\ \vdots & \vdots & \ddots & \vdots \\ \alpha & \alpha & \dots & 1 \end{pmatrix}_{n_i \times n_i} . \tag{7}$$

In our analysis of the OOA data, different families were allowed to have different correlation structures. However, all the structures were **mixed** correlation structures (MCS), which we define as structures that may vary between families but share correlation parameters, so that the parameters for family i take value in $\{\gamma, \rho_1, \rho_2, \alpha\}$. See Chaganty and Deng (2007) for a discussion of ranges of measures of association for binary outcomes with familial patterns of association.

3. Extension of Quasi-Least Squares for Mixed Linear Correlation Structures

3.1. **Quasi-Least Squares.** Here, we briefly describe the method of QLS. Stage one of QLS iterates between updating the regression parameter $\boldsymbol{\beta}$ via (i) solution of the GEE estimating equation for β (Liang and Zeger, 1986):

$$\sum_{i=1}^{m} D_i^{\mathrm{T}} A_i^{-1/2} R_i^{-1}(\boldsymbol{\alpha}) A_i^{-1/2} (\boldsymbol{Y}_i - U_i(\boldsymbol{\beta})) = 0, \tag{8}$$

where $U_i(\boldsymbol{\beta}) = \mathrm{E}(\boldsymbol{Y}_i)$ and $D_i = \frac{\partial U_i}{\partial \boldsymbol{\beta}}$; and (ii) updating the correlation parameter $\boldsymbol{\alpha}$ by minimizing the generalized error sum of squares

$$Q(\boldsymbol{\beta}, R(\boldsymbol{\alpha})) = \sum_{i=1}^{m} \boldsymbol{z}_i^{\mathrm{T}}(\boldsymbol{\beta}) R_i^{-1}(\boldsymbol{\alpha}) \boldsymbol{z}_i(\boldsymbol{\beta}) \tag{9}$$

with respect $\boldsymbol{\alpha} \in \Omega \subseteq \mathbb{R}^s$, where $\boldsymbol{z}_i(\boldsymbol{\beta}) = A_i^{-1/2}(Y_i - U_i) = (z_{i1}, \dots, z_{in_i})$ are known as the Pearson residuals. In addition, Ω is defined as the feasible region for the correlation structure $(R_i(\boldsymbol{\alpha}))_{1,\dots,m}$, so that $\forall \boldsymbol{\alpha} \in \Omega$ and $\forall i \in \{1, \dots, m\}$, $R_i(\boldsymbol{\alpha})$ is positive definite. Stage one of QLS therefore involves

solving the stage one estimating equation

$$D_G = \frac{\partial}{\partial \boldsymbol{\alpha}} \left\{ \sum_{i=1}^{m} \boldsymbol{z}_i^{\mathrm{T}}(\boldsymbol{\beta}) R_i^{-1}(\boldsymbol{\alpha}) \boldsymbol{z}_i(\boldsymbol{\beta}) \right\} = 0. \tag{10}$$

In general, the solution of (10) is not necessarily the minimizer of (9). However, in Section 3.2 we prove that if the $R_i(\boldsymbol{\alpha})$ are linear for all i, the solution of (10) does indeed minimize the generalized error sum of squares (9); Furthermore, the solution will be unique and feasible almost surely.

The QLS stage one estimates $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ and $\hat{\boldsymbol{\delta}}$ for $\boldsymbol{\alpha}$ are the solutions of (8) and the minimizer of (9), respectively. However, Chaganty (1997) proved that the stage one QLS estimate of $\boldsymbol{\alpha}$ is not consistent. In order to correct the asymptotic bias for the QLS stage one estimates, after convergence in stage one, we next solve the stage two estimating equation that depends on the stage one estimates $\hat{\boldsymbol{\delta}}$ (Chaganty and Shults, 1999), for $\boldsymbol{\alpha}$:

$$\sum_{i=1}^{m} \mathrm{tr} \left[\frac{\partial R_i^{-1}(\boldsymbol{\delta})}{\partial \boldsymbol{\delta}} R_i(\boldsymbol{\alpha}) \right] \Bigg|_{\boldsymbol{\delta}=\hat{\boldsymbol{\delta}}} = 0. \tag{11}$$

Theorem 3.2 of Chaganty and Shults (1999) established that if there is a unique root $\hat{\boldsymbol{\alpha}}$ for equation (11) that is a one to one and continuous function of $\hat{\boldsymbol{\delta}}$ and the structure is correctly specified, then $\hat{\boldsymbol{\alpha}}$ is consistent. We refer to $\hat{\boldsymbol{\alpha}}$ as the stage two estimator of $\boldsymbol{\alpha}$, based on which, we obtain the final estimator $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ by again solving the GEE estimating equation (8) for β, evaluated at the stage two estimates for $\boldsymbol{\alpha}$.

3.2. **Results that Justify Application of Quasi-Least Squares for Mixed Linear Correlation Structures.** In this section we first provide general proofs regarding the existence and feasibility of the stage one QLS estimates in section 3.2.1. Next, in section 3.2.2 we prove the consistency of the stage two QLS estimates for the mixed linear correlation structures. The proofs for all results are provided in the appendices.

3.2.1. *General Proof of Feasibility for Stage One QLS Estimates.* We first provide a theorem that establishes the feasibility of the global minimizer for (9).

Theorem 1. *If for each subject i, $R_i(\boldsymbol{\alpha})$ is a differentiable $n_i \times n_i$ matrix, then the global minimizer for (9) in Ω is an inner point of Ω, where Ω is the feasible region of $(R_i(\boldsymbol{\alpha}))_{1,\ldots,m}$.*

Although the stage one QLS estimator of $\boldsymbol{\alpha}$ is not the final estimator, its existence and feasibility is very important because failure to yield feasible estimates in stage one of QLS could cause a breakdown in the first phase of the procedure. For example, Crowder (1995) described the potential for breakdown in iterative procedures such as GEE that can occur when the estimated correlation matrices are not positive definite. Theorem 1 ensures that this type of failure will not occur in stage one of QLS.

However, while Theorem 1 ensures the existence of solutions for the stage one QLS estimating equation (10), it does not guarantee that the root is unique. If (10) has multiple roots, it can be difficult to obtain all the roots. Furthermore, it might not be straightforward to find the global minimizer for (9), if the generalized error sum of squares has several local minimizers. However, for correlation structures that meet the condition in (12), this minimization problem is fairly straightforward, because under this fairly general condition, (9) is convex almost surely, so that there will exist a unique root for (10) almost surely.

Theorem 2. *Suppose each cluster $i \in \{1, \ldots, m\}$ in the data under consideration has correlation structure $R_i(\boldsymbol{\alpha})$. If $\forall\ \boldsymbol{\alpha} \in \Omega$,*

$$\sum_{j=1}^{s} \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_j} c_j = 0 \ if \ and \ only \ if \boldsymbol{c} = 0, \tag{12}$$

then (10) has a unique solution in the feasible region Ω almost surely.

Corollary 3. *Suppose for each cluster $i \in \{1, \ldots, m\}$ of the longitudinal data, we have a linear correlation structure $R_i(\boldsymbol{\alpha})$ of the form (2). Then if $\boldsymbol{\alpha}$ is identifiable, (10) has a unique solution in the feasible region Ω almost surely.*

Theorem 2 provides the criterion (12) that will ensure that the stage one estimating equation (10) has a unique solution; This requirement is fairly general, and is be satisfied by several common structures, including the exchangeable, tri-diagonal (Chaganty and Shults, 1999), BT (Shults et al., 2006), and also by the familial structures implemented in this manuscript.

3.2.2. *Consistency of the Stage Two QLS Estimates for Linear Correlation Structures.* Here we first prove that for linear correlation structures, the stage two estimator exists and is unique, with probability one.

Theorem 4. *If for each cluster $i \in \{1, \ldots, m\}$, the within subject correlation $R_i(\boldsymbol{\alpha})$ has a linear correlation structure of form (2), then the stage two estimating equation (11) has a unique solution with probability one.*

In the proof of Theorem 4 in Appendix A, we provide an explicit solution for the stage two estimating solution for linear correlation structures. Suppose we obtain the stage one estimator $\hat{\boldsymbol{\delta}}$. Define $A_{ij} = R_i^{-1}(\hat{\boldsymbol{\delta}})(R_i(\boldsymbol{e}_j) - R_i(0))$, $M_{jk} = \sum_{i=1}^{m} \mathrm{tr}(A_{ij}A_{ik})$, and $w_j = -\sum_{i=1}^{m} \mathrm{tr}(A_{ij}R_i^{-1}(\hat{\boldsymbol{\delta}})R_i(0))$. Suppose $M = (M_{jk})_{s\times s}$ and $\boldsymbol{w} = (w_1, \ldots, w_s)^{\mathrm{T}}$. We can then express the stage two estimator in a very simple form:

$$\hat{\boldsymbol{\alpha}} = M^{-1}(\hat{\boldsymbol{\delta}})\boldsymbol{w}(\hat{\boldsymbol{\delta}}), \tag{13}$$

which is very helpful with respect to computation, especially when the dimension of $\boldsymbol{\alpha}$ is high. Chaganty and Shults (1999) proved that if a unique solution $\hat{\boldsymbol{\alpha}}$ exists to the stage two estimating equation for $\boldsymbol{\alpha}$ that is a continuous and one to one function of the stage one estimate of $\boldsymbol{\alpha}$, then $\hat{\boldsymbol{\alpha}}$ will

be consistent. Therefore, we have proven that under correct specification of the mixed linear correlation structure, there will exist a unique solution of the stage two estimating equation that will be consistent for $\boldsymbol{\alpha}$.

4. Asymptotic Relative Efficiency Calculations

Here we assess the loss in efficiency that results from incorrectly specifying the mixed correlation structure in the OOA analysis. We assume here that the true structure for cluster i is the mixed structure $R_i(\boldsymbol{\alpha})$ described in Section 5, for $\boldsymbol{\alpha} = (\rho_1, \rho_2, \gamma, \alpha)$, while the working structure is the exchangeable structure $W_i(\gamma) = (1-\gamma)I_{n_i \times n_i} + \gamma J_{n_i \times n_i}$, where $I_{n_i \times n_i}$ is an n_i by n_i identity matrix and $J_{n_i \times n_i}$ is an $n_i \times n_i$ matrix of ones. We consider the exchangeable working structure because this is a popular structure for analysis of clustered data. In addition, we note that the true mixed familial structures include exchangeable structures for OOA families that contain only siblings. Our misidentification scenario therefore represents the situation in which we have correctly assumed that the sibling-sibling correlations are equal, but have incorrectly assumed that the sibling-sibling, sibling-father, sibling-mother, and father-mother correlations are identical.

The efficiencies are calculated using the same approach that was implemented and described in Shults and Morrow (2002) and in Shults et al. (2006). To briefly summarize, we first note that Chaganty (1997) proved that $\sqrt{m}(\hat{\beta} - \beta)$ is asymptotically normal with mean zero and covariance matrix

$$V_w = \lim_{m \to \infty} W_t \left\{ \sum_{i=1}^{m} X_i' A_i^{1/2} W_i^{-1} R_i W_i^{-1} A_i^{1/2} X_i \right\} W_t, \tag{14}$$

where

$$W_t = \left\{ \sum_{i=1}^{m} X_i' A_i^{1/2} W_i^{-1} A_i^{1/2} X_i \right\}^{-1}. \tag{15}$$

If the correct structure was specified, so that $W_i = R_i$, then the covariance matrix V_w can be simplified as $V_t = \lim_{m \to \infty} W_t$.

The efficiency for $\hat{\beta}_j$ was then evaluated as the j^{th} diagonal element of V_t divided by the j^{th} diagonal element of V_w. However, as noted by Sutradhar and Das (1999), $\hat{\gamma}$ may fail to be consistent when the true structure is misspecified, so that the efficiencies should be calculated at the limiting value of $\hat{\gamma}$. We therefore evaluated the efficiencies at $W_i(f(\boldsymbol{\alpha}))$ and $R_i(\boldsymbol{\alpha})$, where $f(\boldsymbol{\alpha})$ is the limiting value of $\hat{\gamma}$ when the mixed correlation structure is misspecified as exchangeable. An algorithm to obtain the limiting value $f(\boldsymbol{\alpha})$ as a function of the true correlation parameter $\boldsymbol{\alpha}$ is provided in the Appendix. Because the efficiencies were calculated as the number of subjects

$m \to \infty$, we assumed that the covariate design for the OOA study was replicated as m increases.

In addition, because the asymptotic distribution for $\hat{\beta}$ is identical for QLS and GEE, the approach for calculation of AREs described in this section also applies when GEE is applied for an exchangeable working structure, but the true structures are mixed familial structures. GEE implements the following moment estimate for the exchangeable structure that is a function of the Pearson residuals z_{ij}:

$$\hat{\alpha}_{GEE} = \frac{\sum_{i=1}^{m} \sum_{k \neq j} z_{ik} z_{ij}}{\sum_{i=1}^{m} \sum_{k=1}^{n_i} (n_i - 1) z_{ik}^2}. \tag{16}$$

It is straightforward to show (Wang and Carey, 2003) that the limiting value of $\hat{\alpha}_{GEE}$ is given by

$$\frac{\sum_{i=1}^{m} \sum_{k \neq j} Corr(y_{ij}, y_{ik})}{\sum_{i=1}^{m} (n_i - 1) n_i} = \frac{\sum_{i=1}^{m} \left(\boldsymbol{e}_i' R_i(\boldsymbol{\alpha}) \boldsymbol{e}_i - n_i \right)}{\sum_{i=1}^{m} (n_i - 1) n_i}, \tag{17}$$

where $\boldsymbol{e}_i$ is an n_i by 1 vector of ones. The limiting values were almost identical for QLS and GEE. As a result, the efficiencies were almost identical for the two approaches.

Table 1 displays the efficiencies for QLS. (An equivalent table for GEE, with almost identical results, is available on request.) Lines 1-3 in Table 1 assess the situation when the father-sibling and sibling-sibling correlations are negligible, but the mother-sibling correlations are non-negligible and get increasingly larger (in going from line 1 to line 3). Lines 4-6 assess the situation when the father-sibling and mother-sibling correlations are negligible, but the sibling-sibling correlations are non-negligible and get increasingly larger (in going from lines 4 to 6). Lines 7-9 assess the situation when the father-sibling and mother-sibling correlations are non-negligible and similar in value, with sibling-sibling correlations that get increasingly larger (in going from lines 7-9). Table 1 indicates that, as we might anticipate, the loss in efficiency is negligible when the true correlations are small, so that the true structure is close to an identity structure, which is a special case of an exchangeable structure (with $\gamma = 0$). However, as the true correlations increase in value, the loss in efficiency can become substantial when the true mixed familial structures are misspecified as exchangeable. For example, as shown in line 6, when $\rho_1 = 0.02$, $\rho_2 = 0.05$, and $\alpha = 0.71$, then the ARE for age is only 79 percent.

The results shown in Table 1 therefore indicate that incorrect application of the exchangeable structure (which is a popular structure in analysis of clustered data) for all families can result in a substantial loss in efficiency in estimation of β. The results in Table 1 are important because it is sometimes claimed that careful modeling of the correlation structure is not crucial, because even if the structure is misspecified, GEE (and QLS) will yield a consistent estimate of the regression parameter. However, our ARE

calculations demonstrate that if the structure is misspecified, even though $\hat{\beta}$ is consistent, we can suffer a substantial loss in efficiency in estimation of β.

TABLE 1. Percent efficiencies for the regression coefficients for the constant term, gender, and age, when the true mixed correlation structure is misspecified as exchangeable. True structure = mixed $R_i(\boldsymbol{\alpha})$ where $\boldsymbol{\alpha} = (\rho_1, \rho_2, \alpha)$; working structure = exchangeable with parameter γ. $limit = f(\boldsymbol{\alpha})$ is the limiting value of $\hat{\gamma}$ when the true mixed structure is misspecified as exchangeable in the analysis of the OOA study.

ρ_1	ρ_2	α	$limit$	constant	gender	age
0.02	0.11	0.05	0.0510	0.99	0.99	0.99
0.02	0.31	0.05	0.0604	0.97	0.98	0.97
0.02	0.41	0.05	0.0652	0.88	0.72	0.88
0.02	0.05	0.41	0.3582	0.94	0.96	0.95
0.02	0.05	0.51	0.4422	0.90	0.92	0.92
0.02	0.05	0.71	0.6092	0.81	0.79	0.79
0.30	0.20	0.50	0.4657	0.95	0.94	0.96
0.30	0.20	0.70	0.6345	0.87	0.83	0.86
0.30	0.20	0.90	0.8029	0.73	0.48	0.53

5. Analysis of the Motivational Study

Here we present our results of the OOA analysis, to demonstrate implementation of the familial structures considered in this manuscript. The OOA population is ideal for studying familial association because the OOA live within a structured and uniform society where most individuals share a common life style. The data considered here represent information on 296 individuals organized from 60 families, of which 33 had both parents and some siblings; 1 had only a father and siblings; 4 had a mother and siblings; and 22 had only siblings. The mean number of siblings in a family was 3.8 (range = 1-11). The mean age was 37.6 (range = 18-85). Recruitment and data collection of the parent study which provided the data for our secondary analysis has been described elsewhere (Wojciechowski et al., 2009).

The main outcome measure used in this analysis was the spherical component of each subject's refractive error. Briefly, refractive error relates to an individual's spectacle prescription. Refractive error is a spherical correction which denotes the power of a spherical lens (a lens whose properties do not change based on orientation) placed in front of a subject's eye to optimize their vision. For some subjects spherical correction alone is sufficient to correct their vision. Lens values for the spherical component of a subject's

refraction can have either a positive or negative value and are expressed in units of optical power called diopters. The outcome for our analysis was the spherical (correction) value, which measures the power of a lens placed in front of the eye that does not depend on orientation. We considered the spherical values of the left eye, right eye, and the average spherical value of both eyes. The covariates we considered included gender ($gender = 1$ for males and $gender = 0$ for females) and the age in years at which the eye exam was conducted.

Our primary objective was to relate the expected spherical values with gender and age. We assumed that the families with both parents and siblings had an extended familial correlation structure (4) with zero correlation between parents, so that $\gamma = 0$ in (4). Families with only a father and siblings, only a mother and siblings, or only siblings, were assumed to have correlation structures (5), (6), and (7), respectively.

Table 2 displays the estimates of the regression parameter estimators. (QLS and GEE share the same asymptotic distribution for $\hat{\beta}$; The results shown here are based on application of a "sandwich based" estimate of the covariance matrix of $\hat{\beta}$ for calculation of standard errors (Chaganty and Shults, 1999), and p-values for the tests that $\beta_j = 0$.) As shown in Table 2, the estimated constant was negative, while the regression coefficients for (male) gender and for age were positive. Although the regression coefficients for age and gender did not differ significantly from zero at a 0.05 level (perhaps as a result of limited power due to the modest number of OOA families studied), the coefficients did differ significantly from zero at a 0.10 level. These results suggest that male gender and higher age are associated with less myopia, where myopia is indicated by negative spherical values.

TABLE 2. The regression parameter estimators for the OOA ophthalmology study. Gender = 1 for male and 0 for female. Age is in years.

	intercept		gender		age	
Outcome	est.	p value	est.	p value	est.	p value
Right Sphere	−2.67	< .0001	0.75	0.067	0.016	0.102
Left Sphere	−2.80	< .0001	0.77	0.051	0.015	0.098
Average Sphere	−2.77	< .0001	0.76	0.055	0.016	0.074

Next, Table 3 displays the QLS estimates of the correlation parameters. Notice that the estimated correlations were similar for the right sphere, left sphere, and average sphere. The estimated correlations were greatest between father and siblings, and smallest between siblings. These findings are consistent with the method of family ascertainment.

TABLE 3. The correlation parameter estimators for the OOA ophthalmology study. $\hat{\rho}_1$ is the estimated correlation between father and siblings, $\hat{\rho}_2$ is the estimated correlation between mother and siblings and $\hat{\alpha}$ is the estimated correlation within siblings.

Outcome	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\alpha}$
Right Sphere	0.2932	0.2241	0.0234
Left Sphere	0.2740	0.1420	0.0130
Average Sphere	0.2880	0.1996	0.0177

6. Discussion

In this paper, we considered QLS, a two-stage approach based on GEE that uses the same estimating equation for estimation of β, but that differs from GEE with respect to estimation of $\boldsymbol{\alpha}$. We proved that the stage one QLS estimates exist and are feasible, while the stage two QLS estimates will be consistent with probability one, for the class of mixed linear correlation structures. We considered familial correlation structures that are members of the class of mixed linear correlation structures. Our general results justified implementation of QLS for the familial structures, in addition to other members of the class of mixed linear structures, e.g. the banded Toeplitz structure that was considered by Shults et al. (2006).

Our work was motivated by a study of spherical optical values in the Old Order Amish (OOA). For this analysis, we implemented QLS for mixed familial correlation structures, which allowed the father-sibling, mother-sibling and sibling-sibling correlations to vary in value. An important feature of the OOA study was that the families varied in size; Our implementation of QLS therefore relaxed the assumption of constant family size and composition that is sometimes made in analysis of familial data.

We also conducted efficiency calculations based on the covariate design of the OOA study, to demonstrate that if the mixed familial structures were the true structures, but were misspecified as exchangeable structures, then we could suffer a serious loss in efficiency in estimation of the regression parameter. Our analysis and efficiency calculations demonstrated that it can be important to carefully model the correlation structure of the data, in order to maximize the information from the data and improve efficiency in estimation of the regression parameter. To encourage the use of the mixed familial correlation structures in practice, we also provide R functions that extend our previous software for application of QLS in R (Xie and Shults, 2009) for implementation of these structures. The R functions, and an R script file that demonstrates their use, is available on request from the first and second authors.

Future research that builds on our methods would be helpful. In extending this exploratory analysis to a larger sample ascertained without regard to myopia status, it will be useful to develop a test that incorporates the gender of each type of family member and to test whether like gender relationships differ from mixed gender relationships within and between families. For example, are the father-son and father-daughter correlations equal in value and are they significantly different from the mother-son and mother-daughter correlations? In addition, in this manuscript we considered spherical values that were measured on the left eye and right eye of each subject, and that were computed as the average of measurements on both eyes. Future work might extend our approach to allow for simultaneous analysis of both eyes. For example, the approach of Shults and Morrow (2002) and Shults, Whitt, and Kumanyika (2004) might be applied to adjust for two sources of correlation: due to the potential similarity of spherical values that are measured on the same subject, or between two members of the same family.

Appendix A. Proofs of Main Results

Proof of Theorem 1. To prove this theorem, we need the following lemma:

Lemma 5. *$R(\boldsymbol{\rho})$ is a differentiable $n \times n$ correlation matrix. Ω_0 is the margin of the feasible region for $R(\boldsymbol{\rho})$. Then we have*

$$\text{Prob}\left(\lim_{\boldsymbol{\rho}\to\Omega_0} z^{\mathrm{T}} R^{-1}(\boldsymbol{\rho}) z = \infty \mid z \in \mathbb{R}^n\right) = 1. \tag{18}$$

We prove Lemma 5 in Appendix B. Here, we directly use this lemma to prove Theorem 1. Suppose the feasible region for $R_i(\boldsymbol{\rho})$ is Ω_i, and the margin of Ω_i is Ω_{i0}. Then the overall feasible region is $\Omega = \cap\Omega_i$, and the margin is $\Omega_0 \subseteqq \cup\Omega_{i0}$. Therefore,

$$\begin{aligned}
&\text{Prob}\left(\lim_{\boldsymbol{\rho}\to\Omega_0} \sum_{i=1}^{m} z_i^{\mathrm{T}} R_i^{-1}(\boldsymbol{\rho}) z_i = \infty \mid z_i \in \mathbb{R}^n, \forall i\right) \\
\geq &\text{Prob}\left(\lim_{\boldsymbol{\rho}\to\cup\Omega_{i0}} \sum_{i=1}^{m} z_i^{\mathrm{T}} R_i^{-1}(\boldsymbol{\rho}) z_i = \infty \mid z_i \in \mathbb{R}^n, \forall i\right) \\
\geq &\text{Prob}\left(\lim_{\boldsymbol{\rho}\to\Omega_{i'0}} z_{i'}^{\mathrm{T}} R_{i'}^{-1}(\boldsymbol{\rho}) z_{i'} = \infty \mid z_{i'} \in \mathbb{R}^n\right) \\
= &1.
\end{aligned} \tag{19}$$

$\Omega = \cup\Omega_i$ is an open set. Because of (19), we know the minimized point of (9) is taken within Ω. And thus the stage one estimators exist and are feasible almost surely.

Proof of Theorem 2. We only need to show that (9) is convex when $\boldsymbol{\alpha} \in \Omega$, and it is equivalent to show that

$$H = \frac{\partial^2 Q(\boldsymbol{\beta}, R(\boldsymbol{\alpha}))}{\partial \boldsymbol{\alpha}^2} \tag{20}$$

is positive definite for all $\boldsymbol{\alpha} \in \Omega$.

Using the fact that

$$\frac{\partial R_i^{-1}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = -R_i^{-1}(\boldsymbol{\alpha}) \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} R_i^{-1}(\boldsymbol{\alpha}), \tag{21}$$

we get

$$H_{jk} = \frac{\partial^2 Q(\boldsymbol{\beta}, R(\boldsymbol{\alpha}))}{\partial \alpha_j \partial \alpha_k} = \sum_{i=1}^{m} \boldsymbol{z}_i^{\mathrm{T}} R_i^{-1}(\boldsymbol{\alpha}) \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_j} R_i^{-1}(\boldsymbol{\alpha}) \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_k} R_i^{-1}(\boldsymbol{\alpha}) \boldsymbol{z}_i \tag{22}$$

Therefore, $\forall$ nonzero $\boldsymbol{x} = (x_1, \ldots, x_s) \in \mathbb{R}^s$,

$$\begin{aligned} \boldsymbol{x}^{\mathrm{T}} H \boldsymbol{x} &= \sum_{j,k} x_j H_{jk} x_k \\ &= \sum_{i=1}^{m} \sum_{j,k} x_j \boldsymbol{z}_i^{\mathrm{T}} R_i^{-1}(\boldsymbol{\alpha}) \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_j} R_i^{-1}(\boldsymbol{\alpha}) \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_k} R_i^{-1}(\boldsymbol{\alpha}) z_i x_k) \end{aligned} \tag{23}$$

Define $\boldsymbol{\gamma}_j^{(i)} = \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_j} R_i^{-1}(\boldsymbol{\alpha}) \boldsymbol{z}_i x_j$ and $G^{(i)} = (\beta_1^{(i)}, \ldots, \beta_s^{(i)})$. Then,

$$\boldsymbol{x}^{\mathrm{T}} H \boldsymbol{x} = \sum_{i=1}^{m} \mathbf{1}^{\mathrm{T}} G^{(i)\mathrm{T}} R_i^{-1}(\boldsymbol{\alpha}) G^{(i)} \mathbf{1} \tag{24}$$

For all $\boldsymbol{\alpha} \in \Omega$, since $R^{-1}(\boldsymbol{\alpha})$ is positive definite, to show $\boldsymbol{x}^{\mathrm{T}} H \boldsymbol{x} > 0$, we only need to show $G^{(i)} \mathbf{1} \neq 0$ when $\boldsymbol{x} \neq 0$.

$$\begin{aligned} G^{(i)} \mathbf{1} &= \sum_{j=1}^{s} \boldsymbol{\gamma}_j^{(i)} \\ &= \left[\sum_{j=1}^{s} \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_j} x_j \right] R_i^{-1}(\boldsymbol{\alpha}) z_i \end{aligned} \tag{25}$$

By hypothesis, for all $\boldsymbol{x} \neq 0$,

$$\left[\sum_{j=1}^{s} \frac{\partial R_i(\boldsymbol{\alpha})}{\partial \alpha_j} x_j \right] \neq 0. \tag{26}$$

Since $\boldsymbol{z}_i \in \mathbb{R}_{n_i}$, $R_i^{-1}(\boldsymbol{\alpha}) z_i$ does not lie in the solution space for (26) almost surely. And therefore, (25) does not equal to 0 almost surely.

Proof of Corollary 3. It is easy to show that for linear correlation structure, $\boldsymbol{\alpha}$ is identifiable if and only if (12) is satisfied.

Proof of Theorem 4. If $R_i(\boldsymbol{\alpha})$ has the form as (2), then

$$\left. \frac{\mathrm{d} R_i^{-1}(\boldsymbol{\delta})}{\mathrm{d} \delta_j} \right|_{\delta = \hat{\delta}} = -R_i^{-1}(\hat{\boldsymbol{\delta}}) (R_i(\boldsymbol{e}_j) - R_i(0)) R_i^{-1}(\hat{\boldsymbol{\delta}}). \tag{27}$$

Plug (2) and (27) into (11) and define $A_{ij} = R_i^{-1}(\hat{\boldsymbol{\delta}})(R_i(\boldsymbol{e}_j) - R_i(0))$, we can rewrite the stage two estimating equation as

$$\sum_{i=1}^{m}\sum_{k=1}^{s} \operatorname{tr}(A_{ij}A_{ik})\alpha_k = -\operatorname{tr}(A_{ij}R_i^{-1}(\hat{\boldsymbol{\delta}})R_i(0)),\ \forall j = 1,\ldots,s. \tag{28}$$

Let $M_{jk} = \sum_{i=1}^{m} \operatorname{tr}(A_{ij}A_{ik})$, $w_j = -\sum_{i=1}^{m} \operatorname{tr}(A_{ij}R_i^{-1}(\hat{\boldsymbol{\delta}})R_i(0))$. Suppose $M = (M_{jk})_{s\times s}$ and $\boldsymbol{w} = (w_1,\ldots,w_s)^{\mathrm{T}}$, then (28) can be written as a linear form

$$M\boldsymbol{\alpha} = \boldsymbol{w}. \tag{29}$$

We have the following lemma, which will be proved in Appendix B.

Lemma 6. *Let $A_{ij} = R_i^{-1}(\hat{\delta})(R_i(e_j) - R_i(0))$, $M_{jk} = \sum_{i=1}^{m} \operatorname{tr}(A_{ij}A_{ik})$. If for each cluster $i \in \{1,\ldots,m\}$, R_i has the linear correlation structure form (2), then $M = (M_{jk})_{s\times s}$ is positive definite.*

Therefore, the stage two estimator $\hat{\boldsymbol{\alpha}} = M^{-1}\boldsymbol{w}$ always exists and is unique.

Appendix B. Proofs of Other Results

Proof of Lemma 5. Suppose eigenvalue, eigenvector pair of $R(\boldsymbol{\rho})$ is

$$\{(\lambda_1(\boldsymbol{\rho}),\ v_1(\boldsymbol{\rho})),\ldots,(\lambda_n(\boldsymbol{\rho}),\ v_n(\boldsymbol{\rho}))\}.$$

Note that the corresponding eigenvalue and eigenvector pairs of $R^{-1}(\boldsymbol{\rho})$ is

$$\{(1/\lambda_1(\boldsymbol{\rho}),\ v_1(\boldsymbol{\rho})),\ldots,(1/\lambda_n(\boldsymbol{\rho}),\ v_n(\boldsymbol{\rho}))\}.$$

Let $\mathcal{X}_1(\boldsymbol{\rho}) = \operatorname{span}\{v_i(\boldsymbol{\rho}) : \lambda_i(\boldsymbol{\rho}) = 0\}$. $\mathcal{X}_2 = \mathbb{R}^n \backslash \mathcal{X}_1$.

Note that the feasible region Ω, which requires all the eigenvalues of R is positive definite, is an open region. It is obvious that Ω, R^{-1} is continuous and differentiable too, since $R^{-1} = \det(R)R^*$, where R^* is the companion matrix of R.

Forall $\boldsymbol{z}$ and M_1, let's fix them temporarily. We take a point $\boldsymbol{\rho}_0$ in the feasible region. $\forall \boldsymbol{\rho}_1 \in \Omega_0$, if $R(\boldsymbol{\rho}_1) = 0$ (I will prove the other situation later), we choose $0 < \epsilon < \|\boldsymbol{z}\|^2 M_1$, $\exists \delta > 0$, such that if $\|\boldsymbol{\rho} - \boldsymbol{\rho}_1\| < \delta$, $\|R(\boldsymbol{\rho}) - R(\boldsymbol{\rho}_1)\|_F < \epsilon$. According to Hoffman-Wielandt Theorem, there exists a permutation $\pi(1),\ \pi(2),\ldots,\pi(n)$ of $1,\ 2,\ldots,n$, such that $\forall \boldsymbol{\rho} \in \Theta_1$,

$$\left(\sum_{i=1}^{n} |\lambda(\boldsymbol{\rho})_{\pi(i)} - \lambda(\boldsymbol{\rho}_1)_i|^2\right)^{\frac{1}{2}} < \|R(\boldsymbol{\rho}) - R(\boldsymbol{\rho}_1)\|_F < \epsilon. \tag{30}$$

From (30), we know that $\forall \boldsymbol{\rho} \in \Theta_1$, $\lambda(\boldsymbol{\rho}) < \epsilon$, and therefore $\frac{1}{\lambda(\boldsymbol{\rho})} > \frac{1}{\epsilon}$. Thus, we have

$$\boldsymbol{z}' R^{-1}(\boldsymbol{\rho})\boldsymbol{z} > \|\boldsymbol{z}\|^2 min\,(1/\lambda\boldsymbol{\rho}) > \|\boldsymbol{z}\|^2/\epsilon > M_1. \tag{31}$$

If $R(\boldsymbol{\rho}_1) \neq 0$, let's suppose $\lambda_1(\boldsymbol{\rho}_1) = \ldots = \lambda_k(\boldsymbol{\rho}_1) = 0$, and $0 < \lambda_{k+1}(\boldsymbol{\rho}_1) \leq \ldots \lambda_n(\boldsymbol{\rho}_1)$. Since $\boldsymbol{\rho}_1 \in \Omega_0$ and $R(\boldsymbol{\rho}_1) \neq 0$, $1 \leq k \leq n-1$. Then $\mathcal{X}_1(\boldsymbol{\rho}_1) = \text{span}\{v_1(\boldsymbol{\rho}), \ldots, v_k(\boldsymbol{\rho})\}$. Obviously, $\mathcal{X}_1 \perp \mathcal{X}_2$. Since $\mathcal{X}_1(\boldsymbol{\rho}_1) \neq \phi$,

$$\text{Prob}\{\text{Proj}(\boldsymbol{z} \mid \mathcal{X}_1) = 0\} = 1. \tag{32}$$

Therefore, with probability 1, $M_2 = \text{Proj}(\boldsymbol{z} \mid \mathcal{X}_1) > 0$.

Suppose $M_3 = \lambda_{k+1}(\boldsymbol{\rho}_1).\forall 0 < \epsilon < \min\{M_2 M_3/4, M_2/(2M_1)\}$, $\exists \delta > 0$, when $\|\boldsymbol{\rho} - \boldsymbol{\rho}_1\| < \delta$, $\|R(\boldsymbol{\rho}) - R(\boldsymbol{\rho}_1)\|_2 < \epsilon$, and $\|R(\boldsymbol{\rho}) - R(\boldsymbol{\rho}_1)\|_F < \epsilon$. From Hoffman-Weilandt Inequality, we know that $\lambda_i(\boldsymbol{\rho}) < \epsilon$, $\forall i = 1, \ldots, k$. (The induction is the same as (30)). According to Stewart Inequality, since $\|R(\boldsymbol{\rho}) - R(\boldsymbol{\rho}_1)\|_2 < \epsilon$, $\text{dist}(\mathcal{X}_1(\boldsymbol{\rho}), \mathcal{X}_1(\boldsymbol{\rho}_1)) \leq 2\epsilon/M_3 = M_2/2$. If $\text{Proj}(\boldsymbol{z} \mid \mathcal{X}_1(\boldsymbol{\rho}_1)) = M_2 \neq 0$, then $Proj(\boldsymbol{z} \mid (X)_1(\boldsymbol{\rho})) > M_2/2 > 0$. Thus,

$$\boldsymbol{z}' R^{-1}(\boldsymbol{\rho}) \boldsymbol{z} \geq \|\text{Proj}(\boldsymbol{z} \mid \mathcal{X}_1(\boldsymbol{\rho}))\|^2/\epsilon > M_1. \tag{33}$$

Therefore, we have

$$\text{Prob}\{\boldsymbol{z}' R^{-1}(\boldsymbol{\rho}) \boldsymbol{z} > M_1\} = \text{Prob}\{\text{Proj}(\boldsymbol{z} \mid \mathcal{X}_1(\boldsymbol{\rho}_1))\} = 1, \ \forall \|\boldsymbol{\rho} - \boldsymbol{\rho}_1\| < \delta. \tag{34}$$

Since Ω_0 is a close region, there exists finite round discs which can cover Ω_0. Within every disc, (34) stands. Therefore, within all the finite round discs, (34) stands. Thus, we proved Lemma 5, and therefore demonstrated that the stage one estimates will have feasible solution with probability 1 for any correlation structure.

Proof of Lemma 6. $\forall x \in \mathbb{R}^s$, we will show $x^{\mathrm{T}} M x > 0$. Suppose $x = (x_1, \ldots, x_s)$.

$$\begin{aligned} x^{\mathrm{T}} M x &= \sum_{i=1}^{m} \sum_{j=1}^{s} \sum_{k=1}^{s} x_j \,\text{tr}(A_{ij} A_{ik}) x_k \\ &= \sum_{i=1}^{m} \sum_{j=1}^{s} \sum_{k=1}^{s} \text{tr}(B_{ij} B_{ik}) \\ &= \sum_{i=1}^{m} \text{tr}\left(B_i^2\right), \end{aligned} \tag{35}$$

where

$$B_{ij} = x_j A_{ij} = R_i^{-1}(\hat{\delta})(R_i(x_j e_j) - R_i(0)), \tag{36}$$

$$\begin{aligned} B_i = \sum_{j=1}^{k} B_{ij} &= R_i^{-1}(\hat{\delta}) \left(\sum_{j=1}^{s} (R_i(x_j e_j) - R_i(0)) \right) \\ &= R_i^{-1}(\hat{\delta})(R_i(x) - R_i(0)). \end{aligned} \tag{37}$$

Thus,

$$B_i^2 = G_i H_i, \tag{38}$$

where

$$G_i = R_i^{-1}(\hat{\delta}) \tag{39}$$

$$H_i = (R_i(x) - R_i(0))R_i^{-1}(\hat{\delta})(R_i(x) - R_i(0)). \tag{40}$$

Since $\hat{\delta}$ is the final stage one estimates for the correlation parameters, by Theorem 1 G_i is positive definite. $\forall y \in \mathbb{R}^s$, $y^{\mathrm{T}}H_i y = [(R_i(x)-R_i(0))y]^{\mathrm{T}}G_i[(R_i(x)-R_i(0))y] \geq 0$, and thus H_i is semi-positive definite.

Kleinman and Athans (1968), in the context of design of suboptimal control systems, obtained that, for any two semi-positive definite matrix A and B,

$$\lambda_n(A)\,\mathrm{tr}(B) \leq \mathrm{tr}(AB) \leq \lambda_1(A)\,\mathrm{tr}(B), \tag{41}$$

where $\lambda_i(A)$ is the ith largest eigenvalue of A.

Because G_i is positive definite, $\lambda_n(G_i) > 0$; and because H_i is semi-positive definite and $H_i \neq 0$, $\mathrm{tr}(H_i) > 0$. Therefore,

$$\mathrm{tr}(B_i^2) = \mathrm{tr}(G_i H_i) \geq \lambda_n(G_i)\,\mathrm{tr}(H_i) > 0. \tag{42}$$

As a result,

$$x^{\mathrm{T}}Mx = \sum_{i=1}^{m} \mathrm{tr}\left(B_i^2\right) > 0, \tag{43}$$

and thus M is positive definite.

Appendix C. The limiting value of the QLS estimate of γ when the true mixed correlation structure is misspecified as exchangeable

Assume the true mixed correlation structures $R_i(\boldsymbol{\alpha})$ have been misspecified as exchangeable $W_i(\gamma)$. Next, using arguments similar to those given in Theorem 3.2 of Chaganty and Shults (1999), we note that $E(Z_i(\beta)Z_i^{'}(\beta)) = \phi R_i(\boldsymbol{\alpha})$. It is then easy to show that the solution to the stage one estimating equation (10) converges in probability to the solution (for γ) to the following estimating equation:

$$trace\left(\sum_{i=1}^{m} \frac{\partial}{\partial \boldsymbol{\alpha}} W_i^{-1}(\gamma) R_i(\boldsymbol{\alpha})\right) = 0. \tag{44}$$

The inverse of an exchangeable structure $W_i(\gamma)$ can be expressed as $W_i^{-1}(\gamma) = \dfrac{1}{(1-\gamma)} I_{n_i} - \dfrac{\gamma}{(1-\gamma)(1+(n_i-1)\gamma)}\,\boldsymbol{e}_j\,\boldsymbol{e}_j'$, where I_{n_i} is the identity matrix and $\boldsymbol{e}_j$ is a $n_i \times 1$ column vector of ones. Next, if we note that $trace(\boldsymbol{e}_j\,\boldsymbol{e}_j' R_i(\boldsymbol{\alpha})) = trace(\boldsymbol{e}_j' R_i(\boldsymbol{\alpha})\boldsymbol{e}_j) = \boldsymbol{e}_j' R_i(\boldsymbol{\alpha}) e_\rangle$, equation 44 can easily be simplified as follows:

$$\sum_{i=1}^{m} n_i - \sum_{i=1}^{m} \frac{1+\gamma^2(n_i-1)}{(1+\gamma(n_i-1))^2} \boldsymbol{e}_j' R_i(\boldsymbol{\alpha}) \boldsymbol{e}_j = 0. \tag{45}$$

In general, a solution $g(\boldsymbol{\alpha})$ (for γ) to (45) can be obtained using the bisection method. We next note that under an assumption of an exchangeable structure, the stage two estimate is obtained as the solution $f(\gamma)$ to the stage two estimating equation (11) that is evaluated at $\hat{\gamma}$ for exchangeable structures $R_i(\gamma)$. Since $\hat{\gamma} \xrightarrow{p} g(\boldsymbol{\alpha})$, it then follows that the limiting value of the stage two estimate for γ converges in probability to $f(g(\boldsymbol{\alpha}))$, so that the limiting value of $\hat{\gamma}$ can be obtained by solving (11) at $\hat{\delta} = g(\boldsymbol{\alpha})$. The stage two estimating equation has a closed form solution for the exchangeable structure that is provided in $(C.3)$ of Shults and Morrow (2002), for $s_i = n_i$ and when $(C.3)$ is calculated over all i, i.e. when $g_i = 1$ for all i, so that we only have one group of subjects.

An algorithm to obtain the limiting value can then be expressed as follows:

(i) For assumed true values of $\boldsymbol{\alpha}$, use the bisection method to obtain a solution $g(\boldsymbol{\alpha})$ to 45.

(ii) Next, obtain the limiting value of $\hat{\gamma}$ by evaluating $(C.3)$ of Shults and Morrow (2002) at $\hat{\tau}_1 = g(\boldsymbol{\alpha})$, where $s_i = n_i$ and $g_i = 1$ for all i.

References

Chaganty, N.R. (1997). An alternative approach to the analysis of longitudinal data via generalized estimating equations. *Journal of Statistical Planning and Inference* **63**, 39–54.

Chaganty, N.R. and Deng, Y. (2007). Ranges of measures of association for familial binary variables. *Communications*

Chaganty, N.R. and Shults, J. (1999). On eliminating the asymptotic bias in the quasi-least squares estimate of the correlation parameter. *Journal of Statistical Planning and Inference* **76**, 127–144.

Crowder, M. (1995). On the use of a working correlation matrix in using generalised linear models for repeated measures. *Biometrika*, **82**, 407-410.

Gleseer, L.J. (1992). A note on the analysis of familial data. *Biometrika* **79(2)**, 412-415

Karlin, S. Cameron, E.C., and P T Williams. (1981). Sibling and parent-offspring correlation estimation with variable family size. *PNAS* **78**, 2664-2668

%QLS SAS macro: A SAS macro for analysis of longitudinal data using quasi-least squares. *UPenn Biostatistics Working Papers.* Working Paper 27. `http://biostats.bepress.com/upennbiostat/papers/art27` This paper is also in press at the *Journal of Statistical Software*

Liang K.Y., Zeger S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Ratcliffe, S. and Shults, J. (2008). GEEQBOX: A MATLAB toolbox for implementation of quasi-least squares and generalized estimating equations. *Journal of Statistical Software* **25(14)**, 1–13.

Shults, J. (1996) *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.

Shults, J. and Chaganty, NR. (1998). Analysis of serially correlated data using

quasi-least squares. *Biometrics* **54**, 1622-1630.
Shults, J., Mazurick, C., and Landis, J.R. (2006), Analysis of repeated bouts of measurements in the framework of generalized estimating equations. *Statistics in Medicine* **25**, 4114-4128.
Shults J. and Morrow, A. (2002). Use of quasi- least squares to adjust for two levels of correlation. *Biometrics* **58**, 521-530. Shults J, Ratcliffe S, Leonard M. (2007). Improved generalized estimating equation analysis via xtqls for implementation of quasi-least squares in Stata. *Stata Journal* **7**, 147-166.
Shults, J., Whitt, C.M. & Kumanyika, S. (2004) Analysis of data with multiple sources of correlation in the framework of generalized estimating equations. *Statistics in Medicine* **23**(20): 3209–3226.
Sutradhar, B.D. & Das, K. (1999). On the efficiency of regression estimators in generalised linear models for longitudinal data. *Biometrika*, **86**, 459-465.
Wang, Y.G. and Carey, V.J. (2003). Working correlation structure misspecification, estimation and covariate design: Implications for generalised estimating equations performance. *Biometrika* **90**, 29–41.
Wojciechowski R., Stambolian, D., Ciner E., Ibay G., Holmes T., Bailey-Wilson J. (2009). Genomewide linkage scans for ocular refraction and meta-analysis of four populations in the Myopia Family Study. *Journal Investigative ophthalmology and Visual Science* **50**, 2024–2032.
Xie, J and Shults, S. (2009). Implementation of quasi-least squares with the R package qlspack. *UPenn Biostatistics Working Papers. Working Paper 32.* `http://biostats.bepress.com/upennbiostat/papers/art32` This paper is under revision at the *Journal of Statistical Software*