Web Appendix for
"Power and sample size requirements for GEE analyses of cluster randomized crossover trials"

Fan Li*, Andrew B. Forbes, Elizabeth L. Turner, John S. Preisser

***frank.li@duke.edu**

A. THE EIGENVALUES OF CORRELATION MATRIX

Recall the two-period nested exchangeable correlation structure is written as

$$R_i(\alpha) = (1-\alpha_0)I_{m_i} + (\alpha_0 - \alpha_1) \oplus_{j=1}^{2} J_{m_{ij}} + \alpha_1 J_{m_i}.$$

For example, when the cluster-period size $m_{i1} = m_{i2} = 3$, the explicit form of R_i is (each block represents a cluster-period)

$$\left(\begin{array}{ccc|ccc} 1 & \alpha_0 & \alpha_0 & \alpha_1 & \alpha_1 & \alpha_1 \\ \alpha_0 & 1 & \alpha_0 & \alpha_1 & \alpha_1 & \alpha_1 \\ \alpha_0 & \alpha_0 & 1 & \alpha_1 & \alpha_1 & \alpha_1 \\ \hline \alpha_1 & \alpha_1 & \alpha_1 & 1 & \alpha_0 & \alpha_0 \\ \alpha_1 & \alpha_1 & \alpha_1 & \alpha_0 & 1 & \alpha_0 \\ \alpha_1 & \alpha_1 & \alpha_1 & \alpha_0 & \alpha_0 & 1 \end{array}\right)$$

Next, to derive the eigenvalues of R_i, we write the left hand side of the characteristic equation as

$$R_i(\alpha) - \lambda I_{m_i} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} (1-\alpha_0-\lambda)I_1 + \alpha_0 J_1 & \alpha_1 J_{12} \\ \alpha_1 J_{21} & (1-\alpha_0-\lambda)I_2 + \alpha_0 J_2 \end{bmatrix},$$

where we denote $I_1 = I_{m_{i1}}$, $I_2 = I_{m_{i2}}$, $J_1 = J_{m_{i1}}$, $J_{12} = J_{21}' = J_{m_{i1}\times m_{i2}}$ and $J_2 = J_{m_{i2}}$ for notational convenience. It is straightforward to see that when $\lambda = 1-\alpha_0$, $R_i(\alpha) - \lambda I_{m_i}$ is less than full rank and has zero determinant. Therefore $\lambda_{i1} = 1-\alpha_0$ is an eigenvalue of R_i. To obtain the rest of the eigenvalues, we now suppose that A is invertible. By Theorem 8.3.4 and 8.4.4 in Graybill[1], any $u \times u$ exchangeable matrix $E = xI + yJ$ is invertible if and only if $x \neq 0$ and $x + uy \neq 0$. Further the inverse can be written as

$$E^{-1} = \frac{1}{x}I - \frac{y}{x(x+uy)}J \tag{1}$$

and the determinant is

$$\det(E) = x^{u-1}(x+uy). \tag{2}$$

This tells us that a sufficient and necessary condition for the existence of A^{-1} is that $\lambda \neq 1 - \alpha_0$ and $\lambda \neq 1 + (m_{i1} - 1)\alpha_0$. By Theorem 8.2.1 in Graybill[1], the determinant of this block matrix is $\det(\boldsymbol{R}_i(\alpha) - \lambda \boldsymbol{I}_{m_i}) = \det(A)\det(\boldsymbol{D} - \boldsymbol{C}A^{-1}\boldsymbol{B})$. This suggests additional eigenvalues are the solutions to $\det(\boldsymbol{D} - \boldsymbol{C}A^{-1}\boldsymbol{B}) = 0$.

Since A is assumed nonsingular, the inverse can be computed by formula (1)

$$A^{-1} = \frac{1}{1 - \alpha_0 - \lambda}\boldsymbol{I}_1 - \frac{\alpha_0}{(1 - \alpha_0 - \lambda)[1 + (m_{i1} - 1)\alpha_0 - \lambda]}\boldsymbol{J}_1.$$

Routine matrix algebra gives

$$\boldsymbol{D} - \boldsymbol{C}A^{-1}\boldsymbol{B} = (1 - \alpha_0 - \lambda)\boldsymbol{I}_2 + \frac{(1 - \alpha_0 - \lambda)\alpha_0 + m_{i1}(\alpha_0^2 - \alpha_1^2)}{1 + (m_{i1} - 1)\alpha_0 - \lambda}\boldsymbol{J}_2,$$

whose determinant could be obtained by formula (2) as

$$\det(\boldsymbol{D} - \boldsymbol{C}A^{-1}\boldsymbol{B}) = (1 - \alpha_0 - \lambda)^{m_{i2}-1}[1 + (m_{i1} - 1)\alpha_0 - \lambda]^{-1} g(\lambda, m_{i1}, m_{i2}, \alpha_0, \alpha_1),$$

where we define

$$g(\lambda, m_{i1}, m_{i2}, \alpha_0, \alpha_1) = \lambda^2 - [(m_{i1} + m_{i2})\alpha_0 + 2(1 - \alpha_0)]\lambda + m_{i1}m_{i2}(\alpha_0^2 - \alpha_1^2) + (m_{i1} + m_{i2})\alpha_0(1 - \alpha_0) + (1 - \alpha_0)^2.$$

Clearly, the eigenvalues are the solutions to $g(\lambda, m_{i1}, m_{i2}, \alpha_0, \alpha_1) = 0$ with λ as the unknown.

The function g is quadratic in λ and has two real roots since the discriminant $\Delta = (m_{i1} - m_{i2})^2\alpha_0^2 + 4m_{i1}m_{i2}\alpha_1^2 > 0$. Therefore the two roots could be obtained by the quadratic formula as

$$\begin{aligned}
\lambda_{i2} =& 1 + \left(\frac{m_i}{2} - 1\right)\alpha_0 - \left\{\left(\frac{m_{i1} - m_{i2}}{2}\right)^2\alpha_0^2 + m_{i1}m_{i2}\alpha_1^2\right\}^{\frac{1}{2}}, \\
\lambda_{i3} =& 1 + \left(\frac{m_i}{2} - 1\right)\alpha_0 + \left\{\left(\frac{m_{i1} - m_{i2}}{2}\right)^2\alpha_0^2 + m_{i1}m_{i2}\alpha_1^2\right\}^{\frac{1}{2}}.
\end{aligned}$$

Note that each of these two eigenvalues has multiplicity one, and therefore λ_{i1} has multiplicity $m_i - 2$.

For $\boldsymbol{R}_i$ to be positive definite, we require all the eigenvalues to be positive. This requires $\alpha_0 < 1$, and

$$1 + \left(\frac{m_i}{2} - 1\right)\alpha_0 > \sqrt{\left(\frac{m_{i1} - m_{i2}}{2}\right)^2\alpha_0^2 + m_{i1}m_{i2}\alpha_1^2} > 0.$$

Equivalently, this suggests the constraints $-1/(m_i/2 - 1) < \alpha_0 < 1$ and

$$\alpha_1^2 < \left(\frac{1 + (m_{i1} - 1)\alpha_0}{m_{i1}}\right)\left(\frac{1 + (m_{i2} - 1)\alpha_0}{m_{i2}}\right).$$

If the cluster-period sizes are balanced such that $m_{i1} = m_{i2} = m/2$, the roots to $g(\lambda, m/2, m/2, \alpha_0, \alpha_1)$ are simplified as $\lambda_2 = 1 + (m/2 - 1)\alpha_0 - m\alpha_1/2$ and $\lambda_3 = 1 + (m/2 - 1)\alpha_0 + m\alpha_1/2$. In this case, the constraints for $\boldsymbol{R}_i$ to be positive definite further reduce to $-1/(m/2-1) < \alpha_0 < 1$ and

$$-\frac{1+(m/2-1)\alpha_0}{m/2} < \alpha_1 < \frac{1+(m/2-1)\alpha_0}{m/2}.$$

B. THE ANALYTICAL INVERSE OF CORRELATION MATRIX

We will assume R_i is nonsingular and hence the corresponding positive eigenvalue conditions hold. Let $A = (1-\alpha_0)I_{m_i} + (\alpha_0 - \alpha_1) \oplus_{j=1}^2 J_{m_{ij}}$ and $B = \alpha_1 J_{m_i}$. By Henderson and Searle[2], the inverse of R_i is

$$R^{-1} = (A+B)^{-1} = A^{-1} - A^{-1}B(I_{m_i} + A^{-1}B)^{-1}A^{-1}. \quad (3)$$

It is easy to verify that the A^{-1} has similar basis matrices as A and can be expanded as $A^{-1} = xI_{m_i} + \oplus_{j=1}^2 y_{ij}J_{m_{ij}}$. Because

$$I_{m_i} = AA^{-1} = (1-\alpha_0)xI_{m_i} + (\alpha_0-\alpha_1)x \oplus_{j=1}^2 J_{m_{ij}} + (1-\alpha_0) \oplus_{j=1}^2 y_{ij}J_{m_{ij}} + (\alpha_0-\alpha_1) \oplus_{j=1}^2 y_{ij}m_{ij}J_{m_{ij}},$$

we must have

$$x = \frac{1}{1-\alpha_0} \qquad y_{ij} = -\frac{\alpha_0-\alpha_1}{(1-\alpha_0)\psi_{ij}},$$

where $\psi_{ij} = 1 + (m_{ij}-1)\alpha_0 - m_{ij}\alpha_1$. Since

$$A^{-1} = \frac{1}{1-\alpha_0}I_{m_i} - \frac{\alpha_0-\alpha_1}{(1-\alpha_0)} \oplus_{j=1}^2 \frac{1}{\psi_{ij}}J_{m_{ij}},$$

we obtain

$$A^{-1}B = \frac{\alpha_1}{1-\alpha_0}J_{m_i} - \frac{(\alpha_0-\alpha_1)\alpha_1}{(1-\alpha_0)}\Big(\oplus_{j=1}^2 \frac{m_{ij}}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i} = \Big(\oplus_{j=1}^2 \frac{\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i}.$$

Observe that

$$I_{m_i} + A^{-1}B = I_{m_i} + \Big(\oplus_{j=1}^2 \frac{\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i},$$

whose inverse again shares the same basis matrices, and could be expressed by $(I_{m_i} + A^{-1}B)^{-1} = I_{m_i} + (\oplus_{j=1}^2 z_{ij}I_{m_{ij}})J_{m_i}$. Observe that

$$\begin{aligned} I_{m_i} &= (I_{m_i} + A^{-1}B)(I_{m_i} + A^{-1}B)^{-1} \\ &= I_{m_i} + \Big(\oplus_{j=1}^2 z_{ij}I_{m_{ij}}\Big)J_{m_i} + \alpha_1\Big(\oplus_{j=1}^2 \frac{1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i} + \alpha_1\Big(\oplus_{j=1}^2 \frac{1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big(\oplus_{j=1}^2 z_{ij}I_{m_{ij}}\Big)J_{m_i} \\ &= I_{m_i} + C. \end{aligned}$$

If the above equality holds, then the (j,j)th block of C must be

$$0 = z_{ij}J_{m_{ij}} + \frac{\alpha_1}{\psi_{ij}}J_{m_{ij}} + \frac{\alpha_1(m_{i1}z_{i1}+m_{i2}z_{i2})}{\psi_{ij}}J_{m_{ij}},$$

which implies

$$\begin{cases} z_{i1} + \alpha_1/\psi_{i1} + \alpha_1(m_{i1}z_{i1}+m_{i2}z_{i2})/\psi_{i1} = 0 & (4) \\ z_{i2} + \alpha_1/\psi_{i2} + \alpha_1(m_{i1}z_{i1}+m_{i2}z_{i2})/\psi_{i2} = 0 & (5) \end{cases}$$

Note that although the above linear system of equations are found by only looking at the diagonal blocks, it turns out that they are sufficient to ensure $C = 0$. To solve for z_{i1} and z_{i2}, we multiply (4) by m_{i1} and (5) by m_{i2} and add them up to obtain

$$\sum_{j=1}^{2} m_{ij}z_{ij} + \sum_{j=1}^{2}\frac{m_{ij}\alpha_1}{\psi_{ij}} + \Big(\sum_{j=1}^{2}\frac{m_{ij}\alpha_1}{\psi_{ij}}\Big)\Big(\sum_{j=1}^{2} m_{ij}z_{ij}\Big) = 0,$$

which gives

$$\sum_{j=1}^{2} m_{ij}z_{ij} = -\sum_{j=1}^{2}\frac{m_{ij}\alpha_1}{\psi_{ij}}\Big/\Big(1+\sum_{j=1}^{2}\frac{m_{ij}\alpha_1}{\psi_{ij}}\Big).$$

Define

$$\gamma_{ij} = \psi_{ij}\Big(1+\sum_{j=1}^{2}\frac{m_{ij}\alpha_1}{\psi_{ij}}\Big),$$

we then deduce from (4) and (5) that

$$z_{ij} = -\frac{\alpha_1}{\psi_{ij}}\Big(1+\sum_{j=1}^{2} m_{ij}z_{ij}\Big) = -\frac{\alpha_1}{\psi_{ij}}\Big/\Big(1+\sum_{j=1}^{2}\frac{m_{ij}\alpha_1}{\psi_{ij}}\Big) = -\frac{\alpha_1}{\gamma_{ij}}.$$

Then

$$\begin{aligned}(I_{m_i}+A^{-1}B)^{-1}A^{-1} &= \Big[I_{m_i} - \Big(\oplus_{j=1}^{2}\frac{\alpha_1}{\gamma_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big]\Big[\frac{1}{1-\alpha_0}I_{m_i} - \oplus_{j=1}^{2}\frac{\alpha_0-\alpha_1}{(1-\alpha_0)\psi_{ij}}J_{m_{ij}}\Big] \\ &= \frac{1}{1-\alpha_0}I_{m_i} - \oplus_{j=1}^{2}\frac{\alpha_0-\alpha_1}{(1-\alpha_0)\psi_{ij}}J_{m_{ij}} - \Big(\oplus_{j=1}^{2}\frac{\alpha_1}{\gamma_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big(\oplus_{j=1}^{2}\frac{1}{\psi_{ij}}I_{m_{ij}}\Big).\end{aligned}$$

Further, routine calculations show that

$$\begin{aligned}A^{-1}B(I_{m_i}+A^{-1}B)^{-1}A^{-1} &= \Big(\oplus_{j=1}^{2}\frac{\alpha_1}{(1-\alpha_0)\psi_{ij}}I_{m_{ij}}\Big)J_{m_i} - \Big(\oplus_{j=1}^{2}\frac{\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big(\oplus_{j=1}^{2}\frac{m_{ij}(\alpha_0-\alpha_1)}{(1-\alpha_0)\psi_{ij}}I_{m_{ij}}\Big) \\ &\quad -\Big(\oplus_{j=1}^{2}\frac{\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big(\oplus_{j=1}^{2}\frac{\alpha_1}{\gamma_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big(\oplus_{j=1}^{2}\frac{1}{\psi_{ij}}I_{m_{ij}}\Big) \\ &= \Big(\oplus_{j=1}^{2}\frac{\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big(\oplus_{j=1}^{2}\frac{1-(m_{i1}/\gamma_{i1}+m_{i2}/\gamma_{i2})\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big).\end{aligned}$$

The inverse is then given in closed form by

$$R_i^{-1}(\alpha) = \frac{1}{1-\alpha_0}I_{m_i} - \oplus_{j=1}^{2}\frac{\alpha_0-\alpha_1}{\psi_{ij}(1-\alpha_0)}J_{m_{ij}} - \Big(\oplus_{j=1}^{2}\frac{\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big)J_{m_i}\Big(\oplus_{j=1}^{2}\frac{1-(m_{i1}/\gamma_{i1}+m_{i2}/\gamma_{i2})\alpha_1}{\psi_{ij}}I_{m_{ij}}\Big).$$

Notably, when $m_{i1} = m_{i2} = m_i/2 = m/2$, we have $\psi_{ij} = 1+(m/2-1)\alpha_0 - m\alpha_1/2 = \lambda_2$, and $\gamma_{ij} = 1+(m/2-1)\alpha_0+m\alpha_1/2 = \lambda_3$, and the inverse simplifies to

$$R_i^{-1}(\alpha) = \frac{1}{1-\alpha_0} I_m - \frac{\alpha_0 - \alpha_1}{\lambda_2(1-\alpha_0)} I_2 \otimes J_{m/2} - \frac{\alpha_1}{\lambda_2 \lambda_3} J_m,$$

which is the inverse of the nested exchangeable correlation in Teerenstra et al.[3] (after correcting the typo in equation (1) of their manuscript).

C. MATRIX-ADJUSTED ESTIMATING EQUATIONS FOR CORRELATION PARAMETERS

We use the matrix-adjusted estimating equations (MAEE) introduced by Preisser et al.[4] to reduce the finite-sample bias in estimating the correlation parameters. Resuming the notations in Section 2.2, we define the collection of upper triangular elements of $\tilde{R}_i$ as $\tilde{\eta}_i = (\tilde{R}_{i12}, \tilde{R}_{i13}, \dots, \tilde{R}_{i(m_i-1)m_i})'$. Then we write the expectation of these correlation estimates as $\rho_i(\alpha) = E(\tilde{\eta}_i)$ such that $\rho_{ijj'} = \alpha_0$ if $1 \leq j < j' \leq m_{i1}$ or $m_{i1}+1 \leq j < j' \leq m_i$, and $\rho_{ijj'} = \alpha_1$ otherwise. Let $S_i = \partial \rho_i / \partial \alpha'$ and W_i to be a $\binom{m_i}{2}$ by $\binom{m_i}{2}$ diagonal working covariance matrix. The α-estimation equations are specified by

$$\sum_{i=1}^{n} S_i' W_i^{-1} (\tilde{\eta}_i - \rho_i(\alpha)) = 0. \tag{6}$$

A simple choice for the working covariance W_i that preserves the consistency of estimation is the identity matrix[5]. Alternative specification of W_i requires an expression for the higher-order moments of Y_{ijk}. If the outcome Y_{ijk} is binary, the diagonal elements of W_i is provided in Prentice[6] and Lu et al.[7]. If the outcome Y_{ijk} is continuous, the diagonal elements is given by Li et al.[8] as $1+\rho_{ijj'}^2$. Denote the cluster leverage[9] as $H_i = D_i(\sum_{i=1}^n D_i' V_i^{-1} D_i)^{-1} D_i' V_i^{-1}$. Preisser et al.[4] corrected for the finite-sample bias in estimating the correlation parameters by setting $\tilde{R}_{ijj'} = G_{ij\cdot} \hat{R}_{i\cdot j'}$ for $j < j'$, where $G_{ij\cdot}$ is the jth row of $G_i = (I_{m_i} - H_i)^{-1}$ and $\hat{R}_{i\cdot j'}$ is the j'th column of of $\hat{R}_i = r_i(\hat{\mu}_i) r_i'(\hat{\mu}_i)$ with $r_i(\hat{\mu}_i) = (r_{i11}(\hat{\mu}_{i11}), r_{i12}(\hat{\mu}_{i12}), \dots, r_{i2m_{i2}}(\hat{\mu}_{i2m_{i2}}))'$. Joint estimation for model parameters based on the θ-estimating equations (GEE) and α-estimating equations (6) follows the iterative steps outlined in Prentice[6]. Finally, unless the dispersion parameter ϕ is set to be unity, we follow Li et al.[8] and update the dispersion estimate from iteration s to $s+1$ as

$$\hat{\phi}^{(s+1)} = \hat{\phi}^{(s)} \frac{\sum_{i=1}^n \sum_{l=1}^{m_i} \tilde{R}_{ill}}{\sum_{i=1}^n m_i - p},$$

where $\tilde{R}_{ill} = G_{il\cdot} \hat{R}_{i\cdot l}$ for $l = 1, \dots, m_i$ and $p = 3$ is the number of regression parameters in the marginal mean model given in Section 2.1.

D. SAMPLE SIZE METHODS ASSUMING NO PERIOD EFFECT

Assuming no period effect, the marginal mean model reduces to $g(\mu_{ijk}) = \tau + \delta X_{ij}$, where τ is the grand mean of the outcome on the link function scale. Then the design matrix for each cluster is $Z_i = [1_2, X_i] \otimes 1_{m/2}$.

For continuous outcomes with the identity link, since $D_i = Z_i$, we have

$$\Sigma_1 = \frac{1}{\sigma^2}\sum_{i=1}^{n} Z_i' R_i^{-1} Z_i = \frac{1}{\sigma^2}\sum_{i=1}^{n}\begin{pmatrix} 1_2' M 1_2 & 1_2 M X_i \\ X_i' M 1_2' & X_i' M X_i \end{pmatrix},$$

Some algebraic simplification shows that the bottom-right element of $n\Sigma_1^{-1}$ is $\sigma_\delta^2 = 4\lambda_2\sigma^2/m$. We then obtain the sample size formula based on a z-test as

$$n = \frac{4(z_{\epsilon_1/2} + z_{\epsilon_2})^2 \lambda_2 \sigma^2}{m\delta_0^2}, \tag{7}$$

which is identical to the formula provided by Giraudeau et al.[10,11]. Observe that in the absence of the period effect, the sample size formula (7) is free of π, namely the required sample size is not affected by the proportion of clusters assigned to each treatment sequence.

For binary outcomes with the logit link, since there is no period effect, $P_1 = Q_2$ and $P_2 = Q_1$. Recall that $D_i = A_i Z_i$, we have

$$\Sigma_1 = \sum_{i=1}^{n} D_i' A_i^{-1/2} R_i^{-1} A_i^{-1/2} D_i = \sum_{i=1}^{n}\begin{pmatrix} 1_2'\Omega_i^{1/2} M \Omega_i^{1/2} 1_2 & 1_2'\Omega_i^{1/2} M \Omega_i^{1/2} X_i \\ X_i'\Omega_i^{1/2} M \Omega_i^{1/2} 1_2 & X_i'\Omega_i^{1/2} M \Omega_i^{1/2} X_i \end{pmatrix} = \frac{nm}{4\lambda_2\lambda_3}\begin{pmatrix} \Lambda & \xi \\ \xi' & \omega \end{pmatrix}.$$

Some algebraic simplification leads to $\omega = (\lambda_2 + \lambda_3)P_1(1 - P_1)$, and

$$\xi = (\lambda_2 + \lambda_3)P_1(1 - P_1) + (\lambda_2 - \lambda_3)\sqrt{P_1(1 - P_1)P_2(1 - P_2)},$$

and the upper-left element is written as

$$\Lambda = (\lambda_2 + \lambda_3)P_1(1 - P_1) + (\lambda_2 + \lambda_3)P_2(1 - P_2) + 2(\lambda_2 - \lambda_3)\sqrt{P_1(1 - P_1)P_2(1 - P_2)}.$$

Then the bottom-right element of $n\Sigma_1^{-1}$ is $\sigma_\delta^2 = 4\lambda_2\lambda_3\Lambda/[m(\omega\Lambda - \xi^2)]$ by matrix inversion. Notice that $\omega\Lambda - \xi^2 = 4\lambda_2\lambda_3 P_1(1 - P_1)P_2(1 - P_2)$, therefore the sample size formula (based on a z-test) becomes

$$n = \frac{(z_{\epsilon_1/2} + z_{\epsilon_2})^2}{m\delta_0^2}\left\{\frac{\lambda_2 + \lambda_3}{P_1(1 - P_1)} + \frac{\lambda_2 + \lambda_3}{P_2(1 - P_2)} + \frac{2(\lambda_2 - \lambda_3)}{\sqrt{P_1(1 - P_1)P_2(1 - P_2)}}\right\}, \tag{8}$$

which equals to half of the required sample size provided by Preisser et al.[12] for a pretest-posttest cross-sectional design. Here, since $\lambda_2 + \lambda_3 = 2 + (m - 2)\alpha_0$ and $\lambda_2 - \lambda_3 = -m\alpha_1$, the required sample size n is increasing in α_0 and decreasing in α_1. Further, if the inter-period correlation coefficient $\alpha_1 = 0$, the CRXO design without period effect can be regarded as a parallel design with twice the sample size. In this case, the sample size formula (8) reduces to

$$n = \frac{(z_{\varepsilon_1/2} + z_{\varepsilon_2})^2(\lambda_2 + \lambda_3)}{m\delta_0^2}\left\{\frac{1}{P_1(1-P_1)} + \frac{1}{P_2(1-P_2)}\right\}, \tag{9}$$

which shares a similar form with the formula provided by Shih[13]. On the other hand, when the intervention effect is small such that $P_1(1-P_1) \approx P_2(1-P_2) \approx \sqrt{P_1(1-P_1)P_2(1-P_2)} \approx P^*(1-P^*)$, (8) can be approximated by

$$n \approx \frac{4(z_{\varepsilon_1/2} + z_{\varepsilon_2})^2\lambda_2}{m\delta_0^2 P^*(1-P^*)}. \tag{10}$$

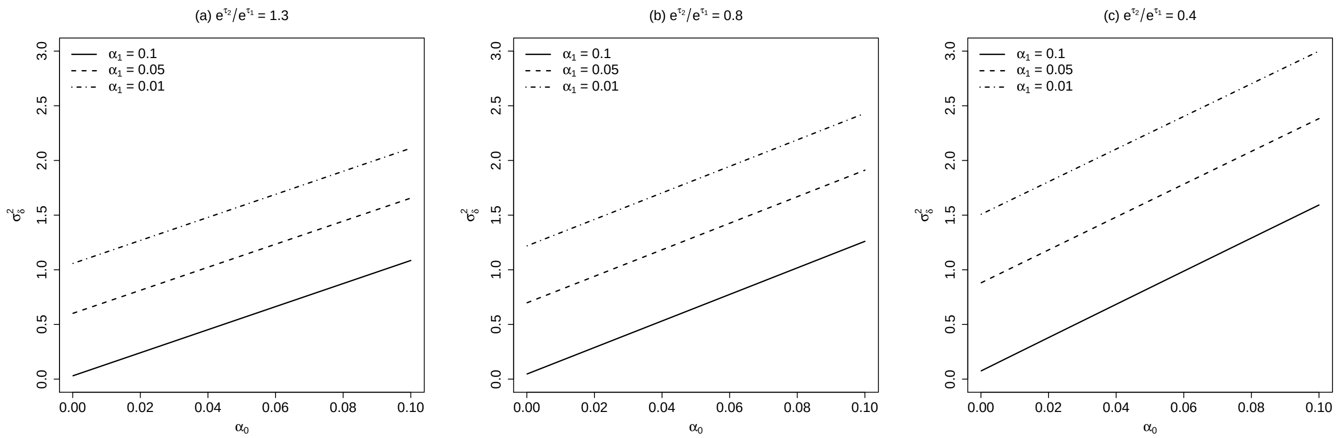
We remark that such an approximation was used by Forbes et al.[14]. Interestingly, formula (10) now shares a similar structure to (7) derived for continuous outcomes assuming no period effect. In this case, λ_2 may be considered as an approximate design effect for binary outcomes.

E. RELATIONSHIP BETWEEN CORRELATIONS AND POWER FOR BINARY OUTCOMES

Although it is challenging to analytically study the relationship between α_0, α_1 and σ_δ^2 with binary outcomes, we have numerically assessed the values of σ_δ^2 as a function of plausible correlation values. We give an illustrative numeric example in Web Figure 1 . The example uses a small cluster size with $m = 20$ so that a wider range of (α_0, α_1) values are plausible (i.e. the resulting $R(\alpha)$ is positive definite). The results for larger value of m are similar and not shown. We choose τ_1 so that a cluster receiving the control condition in the first period has an expected prevalence 0.3. The effect size is chosen by $\delta \approx -0.89$ so that in the absence of period effect (i.e. $\tau_2 = \tau_1$), the expected prevalence will reduce from 0.3 to 0.15 due to the intervention. Note that these numbers mimic the TTANGO trial in Section 5 of the manuscript. Finally, we let the odds ratio $e^{\tau_2}/e^{\tau_1} = 1.3$ or 0.8, representing gently increasing or decreasing period effect in opposite directions. This latter scenario corresponds to larger variance of the intervention effect since the time partially confounds with the intervention for clusters receiving the AB sequence. We further varied $e^{\tau_2}/e^{\tau_1} = 0.4$ to represent an extreme scenario where the period effect $\tau_2 - \tau_1 \approx \delta$. From Web Figure 1 , we observe as the within-period correlation α_0 increases, σ_δ^2 becomes larger and more clusters are required to achieve a fixed power. By contrast, as the inter-period correlation α_1 increases, σ_δ^2 decreases and fewer clusters are required.

F. EXTENSION OF SAMPLE SIZE METHODOLOGY TO ALTERNATIVE LINK FUNCTIONS

For binary outcome Y_{ijk}, we could alternatively specify g as the log function such that the marginal mean of cluster i is $\mu_i = \exp(Z_i\theta)$. We use the Bernoulli model and assume the variance function $h(\mu_{ijk}) = \mu_{ijk}(1-\mu_{ijk})$ and no over-dispersion, so $\phi = 1$. Now define $P_1 = \exp(\tau_1 + \delta_0)$ and $P_2 = \exp(\tau_2)$ to be the expected prevalence in period 1 and 2 for clusters receiving the AB sequence. Similarly, we define $Q_1 = \exp(\tau_1)$ and $Q_2 = \exp(\tau_2 + \delta_0)$ to be the expected prevalence in period 1 and 2 for



Web Figure 1 The variance of the intervention effect σ^2_δ as a function of the within-period correlation α_1 and the inter-period correlation α_2, under different assumptions of period effect.

clusters receiving the BA sequence. Using these quantities, the detectable effect size in terms of the log *risk ratio* (RR) is:

$$\delta_0 = \frac{1}{2}\log(P_1/P_2) - \frac{1}{2}\log(Q_1/Q_2).$$

With the log link, we now have $D_i = \text{diag}(\mu_i)Z_i$, and it follows from Section 3.2 in the main manuscript that

$$\Sigma_1 = \sum_{i=1}^{n} \begin{pmatrix} F_i^{1/2} M F_i^{1/2} & F_i^{1/2} M F_i^{1/2} X_i \\ X_i' F_i^{1/2} M F_i^{1/2} & X_i' F_i^{1/2} M F_i^{1/2} X_i \end{pmatrix} = \frac{nm}{4\lambda_2\lambda_3} \begin{pmatrix} \Lambda & \xi \\ \xi' & \omega \end{pmatrix},$$

where $F_i = \text{diag}\{P_1/(1-P_1), P_2/(1-P_2)\}$ for clusters receiving the AB sequence, $F_i = \text{diag}\{Q_1/(1-Q_1), Q_2/(1-Q_2)\}$ for clusters receiving the BA sequence, and

$$\omega = (\lambda_2 + \lambda_3)[\pi P_1/(1-P_1) + (1-\pi)Q_2/(1-Q_2)],$$

$$\xi = \pi \begin{pmatrix} (\lambda_2 + \lambda_3)P_1/(1-P_1) \\ (\lambda_2 - \lambda_3)\sqrt{P_1/(1-P_1)}\sqrt{P_2/(1-P_2)} \end{pmatrix} + (1-\pi) \begin{pmatrix} (\lambda_2 - \lambda_3)\sqrt{Q_1/(1-Q_1)}\sqrt{Q_2/(1-Q_2)} \\ (\lambda_2 + \lambda_3)Q_2/(1-Q_2) \end{pmatrix},$$

$$\begin{aligned} \Lambda = \ & \pi \begin{pmatrix} (\lambda_2 + \lambda_3)P_1/(1-P_1) & (\lambda_2 - \lambda_3)\sqrt{P_1/(1-P_1)}\sqrt{P_2/(1-P_2)} \\ (\lambda_2 - \lambda_3)\sqrt{P_1/(1-P_1)}\sqrt{P_2/(1-P_2)} & (\lambda_2 + \lambda_3)P_2/(1-P_2) \end{pmatrix} \\ & + (1-\pi) \begin{pmatrix} (\lambda_2 + \lambda_3)Q_1/(1-Q_1) & (\lambda_2 - \lambda_3)\sqrt{Q_1/(1-Q_1)}\sqrt{Q_2/(1-Q_2)} \\ (\lambda_2 - \lambda_3)\sqrt{Q_1/(1-Q_1)}\sqrt{Q_2/(1-Q_2)} & (\lambda_2 + \lambda_3)Q_2/(1-Q_2) \end{pmatrix}. \end{aligned}$$

The revised sample size formula can be obtained once we substitute the new values of δ_0, ω, ξ, Λ in equations (16), (17) of the main paper.

Similar logic extends to the linear probability model, in which g is specified as the identity link. In this case, $P_1 = \tau_1 + \delta_0$, $P_2 = \tau_2$ are the expected prevalence in period 1 and 2 for clusters receiving the AB sequence; $Q_1 = \tau_1$ and $Q_2 = \tau_2 + \delta_0$ are the expected prevalence in period 1 and 2 for clusters receiving the BA sequence; the detectable effect size in *risk difference* (RD) is simply

$$\delta_0 = \frac{1}{2}(P_1 - P_2) - \frac{1}{2}(Q_1 - Q_2).$$

It can be shown that the revised sample size formula is obtained if we specify δ_0 in terms of RD, and substitute ω, ξ, Λ in equations (16), (17) of the main paper with the following:

$$\omega = (\lambda_2 + \lambda_3)[\pi\{P_1(1-P_1)\}^{-1} + (1-\pi)\{Q_2(1-Q_2)\}^{-1}],$$

$$\xi = \pi \begin{pmatrix} (\lambda_2 + \lambda_3)\{P_1(1-P_1)\}^{-1} \\ (\lambda_2 - \lambda_3)\{P_1(1-P_1)P_2(1-P_2)\}^{-1/2} \end{pmatrix} + (1-\pi) \begin{pmatrix} (\lambda_2 - \lambda_3)\{Q_1(1-Q_1)Q_2(1-Q_2)\}^{-1/2} \\ (\lambda_2 + \lambda_3)\{Q_2(1-Q_2)\}^{-1} \end{pmatrix},$$

$$\begin{aligned}\Lambda = \pi &\begin{pmatrix} (\lambda_2 + \lambda_3)\{P_1(1-P_1)\}^{-1} & (\lambda_2 - \lambda_3)\{P_1(1-P_1)P_2(1-P_2)\}^{-1/2} \\ (\lambda_2 - \lambda_3)\{P_1(1-P_1)P_2(1-P_2)\}^{-1/2} & (\lambda_2 + \lambda_3)\{P_2(1-P_2)\}^{-1} \end{pmatrix} \\ + (1-\pi) &\begin{pmatrix} (\lambda_2 + \lambda_3)\{Q_1(1-Q_1)\}^{-1} & (\lambda_2 - \lambda_3)\{Q_1(1-Q_1)Q_2(1-Q_2)\}^{-1/2} \\ (\lambda_2 - \lambda_3)\{Q_1(1-Q_1)Q_2(1-Q_2)\}^{-1/2} & (\lambda_2 + \lambda_3)\{Q_2(1-Q_2)\}^{-1} \end{pmatrix}.\end{aligned}$$

G. EXTENSION OF SAMPLE SIZE METHODOLOGY TO COHORT CRXO TRIALS

Our sample size methodology readily extends to cohort CRXO designs, in which the same set of individuals are included in both periods for a cluster. The GEE analyses of cohort CRXO trials could still be based on the marginal mean model (1) in the main manuscript, but the correlation structure should additionally reflect the association between repeated measurements from the same individual. More specifically, three types of correlations should be considered: the within-period correlation, $\text{corr}(Y_{ijk}, Y_{ijk'}) = \alpha_0$ for $k \neq k'$ and $j = 1, 2$, the inter-period correlation, $\text{corr}(Y_{i1k}, Y_{i2k'}) = \alpha_1$ for $k \neq k'$, and the within-individual correlation, $\text{corr}(Y_{i1k}, Y_{i2k}) = \alpha_2$. Assume $m/2$ individuals are included in each cluster, we follow Li et al.[8] and define the two-period block exchangeable correlation structure for cluster i as

$$R_i(\alpha) = (1 - \alpha_0 + \alpha_1 - \alpha_2)I_m + (\alpha_2 - \alpha_1)J_2 \otimes I_{m/2} + (\alpha_0 - \alpha_1)I_2 \otimes J_{m/2} + \alpha_1 J_m,$$

where $\alpha = (\alpha_0, \alpha_1, \alpha_2)'$ is the vector of correlation parameters, I_u is a u-dimensional identity matrix, $J_s = 1_s 1_s'$ is an s by s matrix of ones. For example, suppose the ith cohort size is $m/2 = 3$, the explicit form of the block exchangeable matrix is

$$\left(\begin{array}{ccc|ccc} 1 & \alpha_0 & \alpha_0 & \alpha_2 & \alpha_2 & \alpha_1 \\ \alpha_0 & 1 & \alpha_0 & \alpha_1 & \alpha_2 & \alpha_1 \\ \alpha_0 & \alpha_0 & 1 & \alpha_1 & \alpha_1 & \alpha_2 \\ \hline \alpha_2 & \alpha_1 & \alpha_1 & 1 & \alpha_0 & \alpha_0 \\ \alpha_1 & \alpha_2 & \alpha_1 & \alpha_0 & 1 & \alpha_0 \\ \alpha_1 & \alpha_1 & \alpha_2 & \alpha_0 & \alpha_0 & 1 \end{array}\right)$$

The properties of the block exchangeable correlation structure has been studied by Li et al.[8] in the context of cohort stepped wedge CRTs and we adapt their results to inform the design of cohort CRXO trials. According to Li et al.[8], $R_i(\alpha)$ has the following four eigenvalues

$$\kappa_1 = 1 - \alpha_0 + \alpha_1 - \alpha_2, \quad \kappa_2 = 1 - \alpha_0 - \alpha_1 + \alpha_2,$$
$$\kappa_3 = 1 + \left(\frac{m}{2} - 1\right)(\alpha_0 - \alpha_1) - \alpha_2, \quad \kappa_4 = 1 + \left(\frac{m}{2} - 1\right)(\alpha_0 + \alpha_1) + \alpha_2.$$

Therefore, valid correlation values are among those such that $R_i(\alpha)$ is positive definite, and can be determined efficiently by directly assessing the positivity of above eigenvalues. Furthermore, the closed-form inverse of the block exchangeable correlation is also available:

$$R_i^{-1}(\alpha) = \frac{1}{\kappa_1} I_m - \frac{\alpha_2 - \alpha_1}{\kappa_1 \kappa_2} J_2 \otimes I_{m/2} - \frac{\alpha_0 - \alpha_1}{\kappa_1 \kappa_3} I_2 \otimes J_{m/2} + \left\{ \frac{(\alpha_2 - \alpha_1)(\alpha_0 - \alpha_1)}{\kappa_1 \kappa_2 \kappa_3} + \frac{\alpha_2 \alpha_0 - \alpha_1}{\kappa_2 \kappa_3 \kappa_4} \right\} J_m.$$

With the block exchangeable correlation matrix, one could repeat the derivation presented in Section 3.1 and 3.2 of the main manuscript and obtain the appropriate sample size formulae for cohort designs. We provide the main idea as follows.

For continuous outcomes with identity link, we have

$$\Sigma_1 = \frac{1}{\sigma^2} \sum_{i=1}^{n} Z_i' R_i^{-1} Z_i = \frac{1}{\sigma^2} \sum_{i=1}^{n} \begin{pmatrix} M & M X_i \\ X_i' M & X_i' M X_i \end{pmatrix},$$

but the constant matrix now becomes

$$M = \frac{m}{4\kappa_3 \kappa_4} \begin{pmatrix} \kappa_3 + \kappa_4 & \kappa_3 - \kappa_4 \\ \kappa_3 - \kappa_4 & \kappa_3 + \kappa_4 \end{pmatrix}.$$

It can then be shown by matrix inversion that the bottom-right element of $n\Sigma_1^{-1}$ is

$$\sigma_\delta^2 = \frac{\kappa_3 \sigma^2}{m\pi(1-\pi)},$$

and the the required total sample size to achieve a prescribed type I error rate ϵ_1 and type II error rate ϵ_2 for a z-test becomes

$$n = (z_{\epsilon_1/2} + z_{\epsilon_2})^2 \frac{\kappa_3 \sigma^2}{\pi(1-\pi) m \delta_0^2},$$

Given the above sample size formula, the new variance inflation factor (design effect) remains to be an eigenvalue of the block exchangeable correlation matrix, κ_3. This variation inflation factor suggests that the required number of clusters increases as the within-period correlation α_0 increases, and as the inter-period correlation α_1 or the within-individual correlation α_2 decreases.

For binary outcomes with the canonical logit link, the revised sample size formula for cohort trials can be obtained by essentially replacing λ_2, λ_3 in equations (16), (17) in the main manuscript with κ_3 and κ_4, respectively.

References

1. Graybill FA. *Matrices with Applications in Statistics*. Wadsworth, Belmont, CA: Brooks/Cole; 1983.
2. Henderson HV, Searle SR. On deriving the inverse of a sum of matrices. *Siam Review.* 1981;23:53–60.
3. Teerenstra S, Lu B, Preisser JS, Van Achterberg T, Borm GF. Sample size considerations for GEE analyses of three-level cluster randomized trials. *Biometrics.* 2010;66(4):1230–1237.
4. Preisser JS, Lu B, Qaqish BF. Finite sample adjustments in estimating equations and covariance estimators for intracluster correlations. *Statistics in Medicine.* 2008;27(27):5764–5785.
5. Sharples K, Breslow N. Regression analysis of correlated binary data: some small sample results for the estimating equation approach. *Journal of Statistical Computation and Simulation.* 1992;42(1-2):1–20.
6. Prentice RL. Correlated binary regression with covariates specific to each binary observation.. *Biometrics.* 1988;44(4):1033–1048.
7. Lu B, Preisser JS, Qaqish BF, Suchindran C, Bangdiwala SI, Wolfson M. A comparison of two bias-corrected covariance estimators for generalized estimating equations. *Biometrics.* 2007;63(3):935–941.
8. Li F, Turner EL, Preisser JS. Sample size determination for GEE analyses of stepped wedge cluster randomized trials. *Biometrics.* 2018;Early view:000-000.
9. Preisser JS, Qaqish BF. Deletion diagnostics for generalised estimating equations. *Biometrika.* 1996;83(3):551–562.
10. Giraudeau B, Ravaud P, Donner A. Sample size calculation for cluster randomized cross-over trials. *Statistics in Medicine.* 2008;27(27):5578–5585.

11. Giraudeau B, Ravaud P, Donner A. Correction to “Sample size calculation for cluster randomized cross-over trials". *Statistics in Medicine.* 2009;28(4):720.

12. Preisser JS, Young ML, Zaccaro DJ, Wolfson M. An integrated population-averaged approach to the design, analysis and sample size determination of cluster-unit trials. *Statistics in Medicine.* 2003;22(8):1235–1254.

13. Shih WJ. Sample size and power calculations for pediodontal and other studies with clustered samples using the method of generalized estimating equations. *Biometrical Journal.* 1997;8:899–908.

14. Forbes AB, Akram M, Pilcher D, Cooper J, Bellomo R. Cluster randomised crossover trials with binary data and unbalanced cluster sizes: Application to studies of near-universal interventions in intensive care. *Clinical Trials.* 2015;12(1):34–44.

H. WEB TABLES

Web Table 1 Summary of parameter constellation and convergence rates for GEE analyses of simulated continuous outcomes.

τ_1	$\tau_2 - \tau_1$	$\delta/\phi^{1/2}$	α_0	α_1	n	m	Convergence (Size)[a]	Convergence (Power)[b]
0	-0.2	-0.40	0.05	0.025	8	90	876	877
0	-0.2	-0.40	0.05	0.025	10	50	988	990
0	-0.2	-0.40	0.07	0.035	12	40	999	999
0	-0.2	-0.40	0.07	0.035	8	140	844	844
0	-0.2	-0.40	0.07	0.035	14	30	998	1000
0	-0.2	-0.30	0.07	0.035	12	150	1000	1000
0	-0.2	-0.30	0.07	0.035	16	60	996	1000
0	-0.2	-0.30	0.10	0.050	14	120	1000	1000
0	-0.2	-0.30	0.10	0.050	18	70	999	1000
0	-0.2	-0.25	0.10	0.050	20	130	1000	1000
0	-0.1	-0.30	0.05	0.040	10	80	977	977
0	-0.1	-0.25	0.05	0.040	12	90	1000	1000
0	-0.1	-0.25	0.07	0.035	16	120	1000	1000
0	-0.1	-0.25	0.07	0.035	18	100	1000	1000
0	-0.1	-0.25	0.07	0.035	16	150	1000	1000
0	-0.1	-0.25	0.10	0.050	24	104	1000	1000
0	-0.1	-0.25	0.10	0.050	26	70	1000	999
0	-0.1	-0.25	0.10	0.050	20	90	1000	1000
0	-0.1	-0.20	0.10	0.080	22	80	999	997
0	-0.1	-0.20	0.10	0.080	18	120	1000	1000

[a] Convergence rates (out of 1000) in simulation scenarios for studying test size.

[b] Convergence rates (out of 1000) in simulation scenarios for studying power.

Web Table 2 Summary of parameter constellation and convergence rates for GEE analyses of simulated binary outcomes.

e_1^τ	e_2^τ/e_1^τ	e^δ	α_0	α_1	n	m	Convergence (Size)[a]	Convergence (Power)[b]
0.5	0.8	0.4	0.05	0.025	8	90	805	783
0.5	0.8	0.4	0.05	0.025	10	36	988	975
0.5	0.8	0.4	0.07	0.035	12	30	985	994
0.5	0.8	0.4	0.07	0.035	8	150	825	798
0.5	0.8	0.4	0.07	0.035	14	24	986	977
0.5	0.8	0.5	0.07	0.035	10	160	998	977
0.5	0.8	0.5	0.07	0.035	12	90	969	971
0.5	0.8	0.5	0.10	0.050	16	50	918	910
0.5	0.8	0.6	0.10	0.050	18	170	1000	1000
0.5	0.8	0.6	0.10	0.050	22	130	997	1000
0.3	0.8	0.4	0.05	0.040	10	50	912	874
0.3	0.8	0.5	0.05	0.040	12	70	949	932
0.3	0.9	0.5	0.07	0.035	14	80	967	966
0.3	0.9	0.5	0.07	0.035	16	100	964	971
0.3	0.9	0.5	0.07	0.035	14	130	998	996
0.3	0.9	0.6	0.10	0.050	24	170	1000	998
0.3	0.9	0.6	0.10	0.050	26	110	1000	1000
0.3	0.9	0.6	0.10	0.080	20	70	929	941
0.3	0.9	0.6	0.10	0.080	18	104	1000	999
0.3	0.9	0.6	0.10	0.080	24	50	905	905

[a] Convergence rates (out of 1000) in simulation scenarios for studying test size.

[b] Convergence rates (out of 1000) in simulation scenarios for studying power.

Web Table 3 Simulation scenarios, predicted power based on z-test and t-test, along with the corresponding empirical power of GEE analyses using different variance estimators for continuous outcomes. MAEE are used for bias-corrected estimation of correlation parameters.

							z-test						t-test					
τ_1	$\tau_2 - \tau_1$	$\delta/\phi^{1/2}$	α_0	α_1	n	m	Pred[a]	MB[b]	BC0[c]	BC1[d]	BC2[e]	BC3[f]	Pred[a]	MB[b]	BC0[c]	BC1[d]	BC2[e]	BC3[f]
0	-0.2	-0.40	0.05	0.025	8	90	0.961	0.951	0.979	0.945	0.902	0.961	0.850	0.853	0.909	0.852	0.745	0.870
0	-0.2	-0.40	0.05	0.025	10	50	0.946	0.947	0.976	0.949	0.911	0.956	0.865	0.881	0.926	0.880	0.819	0.896
0	-0.2	-0.40	0.07	0.035	12	40	0.930	0.913	0.937	0.913	0.880	0.922	0.864	0.861	0.890	0.864	0.801	0.869
0	-0.2	-0.40	0.07	0.035	8	140	0.954	0.936	0.964	0.936	0.893	0.948	0.833	0.833	0.902	0.832	0.724	0.852
0	-0.2	-0.40	0.07	0.035	14	30	0.925	0.920	0.940	0.922	0.899	0.927	0.872	0.890	0.913	0.889	0.855	0.895
0	-0.2	-0.30	0.07	0.035	12	150	0.922	0.893	0.927	0.896	0.865	0.904	0.853	0.831	0.879	0.829	0.783	0.844
0	-0.2	-0.30	0.07	0.035	16	60	0.910	0.890	0.910	0.888	0.859	0.898	0.863	0.845	0.877	0.843	0.812	0.851
0	-0.2	-0.30	0.10	0.050	14	120	0.876	0.868	0.897	0.868	0.823	0.875	0.809	0.799	0.845	0.793	0.747	0.810
0	-0.2	-0.30	0.10	0.050	18	70	0.905	0.904	0.913	0.904	0.886	0.909	0.864	0.873	0.893	0.873	0.837	0.885
0	-0.2	-0.25	0.10	0.050	20	130	0.879	0.867	0.885	0.866	0.828	0.874	0.839	0.816	0.857	0.815	0.787	0.824
0	-0.1	-0.30	0.05	0.040	10	80	0.955	0.950	0.975	0.950	0.924	0.957	0.880	0.908	0.937	0.903	0.842	0.917
0	-0.1	-0.25	0.05	0.040	12	90	0.935	0.928	0.947	0.926	0.900	0.932	0.871	0.881	0.916	0.883	0.829	0.891
0	-0.1	-0.25	0.07	0.035	16	120	0.882	0.884	0.902	0.884	0.855	0.890	0.829	0.828	0.869	0.826	0.790	0.838
0	-0.1	-0.25	0.07	0.035	18	100	0.900	0.879	0.898	0.878	0.859	0.887	0.857	0.847	0.869	0.846	0.809	0.853
0	-0.1	-0.25	0.07	0.035	16	150	0.901	0.898	0.913	0.896	0.869	0.902	0.852	0.862	0.881	0.861	0.823	0.867
0	-0.1	-0.25	0.10	0.050	24	104	0.916	0.914	0.921	0.913	0.903	0.915	0.889	0.899	0.907	0.896	0.878	0.902
0	-0.1	-0.25	0.10	0.050	26	70	0.906	0.905	0.917	0.905	0.891	0.908	0.880	0.888	0.899	0.888	0.875	0.891
0	-0.1	-0.25	0.10	0.050	20	90	0.848	0.843	0.873	0.841	0.819	0.852	0.804	0.813	0.837	0.812	0.785	0.817
0	-0.1	-0.20	0.10	0.080	22	80	0.896	0.891	0.900	0.890	0.875	0.894	0.863	0.869	0.880	0.868	0.841	0.874
0	-0.1	-0.20	0.10	0.080	18	120	0.894	0.889	0.905	0.886	0.872	0.894	0.850	0.863	0.880	0.859	0.815	0.870

[a] Pred: Predicted power.

[b] MB: Model-based variance.

[c] BC0: Uncorrected sandwich variance of Liang and Zeger (1986).

[d] BC1: Bias-corrected sandwich variance of Kauermann and Carroll (2001).

[e] BC2: Bias-corrected sandwich variance of Mancl and DeRouen (2001).

[f] BC3: Bias-corrected sandwich variance of Fay and Graubard (2001).

Web Table 4 Simulation scenarios, predicted power based on z-test and t-test, along with the corresponding empirical power of GEE analyses using different variance estimators for binary outcomes. MAEE are used for bias-corrected estimation of correlation parameters.

							z-test						t-test					
e_1^τ	e_2^τ/e_1^τ	e^δ	α_0	α_1	n	m	Pred[a]	MB[b]	BC0[c]	BC1[d]	BC2[e]	BC3[f]	Pred[a]	MB[b]	BC0[c]	BC1[d]	BC2[e]	BC3[f]
0.5	0.8	0.4	0.05	0.025	8	90	0.978	0.968	0.985	0.967	0.940	0.969	0.890	0.900	0.943	0.902	0.837	0.918
0.5	0.8	0.4	0.05	0.025	10	36	0.928	0.915	0.941	0.914	0.879	0.920	0.838	0.839	0.890	0.840	0.779	0.856
0.5	0.8	0.4	0.07	0.035	12	30	0.919	0.923	0.944	0.920	0.893	0.927	0.849	0.875	0.909	0.871	0.816	0.881
0.5	0.8	0.4	0.07	0.035	8	150	0.975	0.962	0.984	0.962	0.915	0.970	0.882	0.862	0.922	0.863	0.792	0.883
0.5	0.8	0.4	0.07	0.035	14	24	0.920	0.922	0.939	0.922	0.897	0.926	0.866	0.885	0.907	0.882	0.844	0.888
0.5	0.8	0.5	0.07	0.035	10	160	0.930	0.916	0.943	0.915	0.870	0.925	0.840	0.834	0.886	0.836	0.758	0.845
0.5	0.8	0.5	0.07	0.035	12	90	0.931	0.920	0.947	0.923	0.877	0.928	0.866	0.856	0.899	0.850	0.820	0.866
0.5	0.8	0.5	0.10	0.050	16	50	0.892	0.876	0.899	0.874	0.848	0.885	0.841	0.831	0.863	0.826	0.782	0.842
0.5	0.8	0.6	0.10	0.050	18	170	0.858	0.853	0.874	0.853	0.827	0.856	0.808	0.813	0.841	0.811	0.776	0.824
0.5	0.8	0.6	0.10	0.050	22	130	0.904	0.917	0.934	0.918	0.907	0.922	0.872	0.898	0.911	0.898	0.875	0.906
0.3	0.8	0.4	0.05	0.040	10	50	0.941	0.946	0.967	0.941	0.906	0.946	0.858	0.878	0.918	0.875	0.817	0.896
0.3	0.8	0.5	0.05	0.040	12	70	0.938	0.961	0.972	0.961	0.932	0.966	0.877	0.916	0.946	0.914	0.864	0.926
0.3	0.9	0.5	0.07	0.035	14	80	0.870	0.877	0.896	0.871	0.846	0.879	0.803	0.826	0.863	0.826	0.772	0.837
0.3	0.9	0.5	0.07	0.035	16	100	0.930	0.932	0.946	0.926	0.912	0.930	0.888	0.900	0.922	0.901	0.867	0.910
0.3	0.9	0.5	0.07	0.035	14	130	0.918	0.922	0.934	0.915	0.890	0.921	0.863	0.873	0.902	0.871	0.833	0.880
0.3	0.9	0.6	0.10	0.050	24	170	0.857	0.859	0.874	0.854	0.838	0.860	0.822	0.836	0.851	0.832	0.805	0.836
0.3	0.9	0.6	0.10	0.050	26	110	0.853	0.864	0.873	0.858	0.842	0.862	0.822	0.833	0.853	0.837	0.808	0.839
0.3	0.9	0.6	0.10	0.080	20	70	0.886	0.895	0.913	0.891	0.867	0.897	0.847	0.864	0.882	0.860	0.834	0.865
0.3	0.9	0.6	0.10	0.080	18	104	0.913	0.926	0.940	0.921	0.902	0.926	0.873	0.896	0.915	0.890	0.861	0.899
0.3	0.9	0.6	0.10	0.080	24	50	0.881	0.876	0.890	0.875	0.857	0.881	0.849	0.849	0.871	0.850	0.832	0.856

[a] Pred: Predicted power.

[b] MB: Model-based variance.

[c] BC0: Uncorrected sandwich variance of Liang and Zeger (1986).

[d] BC1: Bias-corrected sandwich variance of Kauermann and Carroll (2001).

[e] BC2: Bias-corrected sandwich variance of Mancl and DeRouen (2001).

[f] BC3: Bias-corrected sandwich variance of Fay and Graubard (2001).