Published in final edited form as:

Commun Stat Theory Methods. 2019 ; 48(6): 1435–1448. doi:10.1080/03610926.2018.1433849.

Calculating Power for the General Linear Multivariate Model With One or More Gaussian Covariates

S.M. Kreidler,

Kaztronix

B.M. Ringham,

Department of Biostatistics and Informatics, University of Colorado Denver

K.E. Muller,

Department of Health Outcomes and Policy, University of Florida, Gainesville

D.H. Glueck*

Department of Biostatistics and Informatics, University of Colorado Denver

Abstract

We describe a noncentral χ^2 power approximation for hypotheses about fixed predictors in general linear multivariate models with one or more Gaussian covariates. The results apply to both single and multiple parameter hypotheses. The approach extends power approximations for models with only fixed predictors, and for models with a single Gaussian covariate. The new method approximates the noncentrality parameter under the alternative hypothesis using a Taylor series expansion for the matrix-variate beta distribution of type I. We used a Monte Carlo simulation to evaluate the accuracy of both the novel power approximation, and published power approximations. The simulation study accounted for randomness in both the predictors and the errors. We varied the number of outcomes, the number of parameters in the hypothesis, the per-treatment sample size, and the correlations between the random predictors and the outcomes. We demonstrate that our approximation is more accurate than published methods, both in small and large samples. We show that the run time for a single power calculation with the new method is on the order of milliseconds, compared to an average empirical simulation time of roughly three minutes. Approximate and simulated power can be calculated using the free, open-source rPowerlib package. (<http://github.com/SampleSizeShop/rPowerlib>).

Keywords

Multivariate linear models; power; covariates; oral cancer

1. Introduction

The general linear multivariate model allows for two types of predictors: fixed predictors, such as treatment group, that are set by design, and random covariates, such as weight, that

*Corresponding author: Deborah H. Glueck, 13001 E 17th Pl, Mail Stop B119, Aurora, CO 80045, U.S.A.; deborah.glueck@ucdenver.edu.

are only observed after drawing a sample. We describe a power approximation for tests of hypotheses about fixed predictors in the general linear multivariate model with one or more random Gaussian covariates. The results apply to both single and multiple parameter hypotheses. Increasing the accuracy of power calculations in the presence of covariates ensures that a study will answer the scientific question of interest, while minimizing risks to participants and research costs.

Although the inclusion of Gaussian covariates does not alter data analytic methods, randomness in the predictors changes the distribution theory for power and sample size analysis (Sampson, 1974; Glueck and Muller, 2003). Sampson (1974) described power calculations that depend on the value of the covariates as *conditional*. When random covariates appear in the design matrix, the power becomes a random variable. Sampson defined the *unconditional power* as the expected value of the power across all possible realizations of the design matrix.

Several authors have proposed methods for unconditional power analysis in designs with both fixed predictors and Gaussian predictors. Power and sample size theory is described thoroughly for the general linear multivariate model with only fixed predictors (Muller and Peterson, 1984; Muller and Barton, 1989; Muller and Benignus, 1992; Muller et al., 2007; Sampson, 1974). For multiple covariates, Shieh (2005) proposed an asymptotic method based on the expectation of the design matrix. Glueck and Muller (2003) described a power approximation for the special case of a single Gaussian covariate, based on exact distribution theory for the noncentrality parameter. Glueck and Muller (2003) were able to express the random noncentrality as a weighted sum of independent χ^2 random variables. The Glueck and Muller (2003) approach yielded exact distributional results for the noncentrality parameter because, for the single Gaussian covariate case, the noncentrality parameter is scalar, and thus is equal to its trace (Lemma 1, Glueck and Muller, 2003).

The goal of this work is to generalize the Glueck and Muller (2003) power approximation for models with fixed predictors and a single Gaussian covariate to allow for multiple Gaussian covariates. With two or more Gaussian predictors, however, the special case approach utilized by Glueck and Muller (2003) is no longer applicable. Instead, we estimate the noncentrality parameter using a Taylor series expansion based on the matrix-variate beta distribution of type I. In addition, we propose to modify the numerator degrees of freedom for the distribution of the test statistic under the null to account for the additional column or columns of Gaussian predictors in the design matrix.

We note that power results can always be obtained by simulation. In fact, users who wish to compute empirical power via simulation can do so using the R package associated with this manuscript. The primary drawback to this approach is the time and effort required to run simulations. For the designs considered in this manuscript, the empirical power calculations required several minutes to run, while the corresponding approximate power results could be computed in a few milliseconds. Detailed timing results are shown in Section 4.

We describe a fast, accurate, unconditional power approximation for general linear multivariate models with one or more Gaussian covariates, when the hypothesis involves

only the fixed predictors. The method uses a reduced covariance matrix, scaled hypothesis sum of squares, and adjusted degrees of freedom as inputs for a noncentral $\mathcal{F}$ approximation for power.

The remainder of the manuscript is organized as follows. Section 2 gives notation for the general linear multivariate model, the general linear hypothesis, and the Hotelling-Lawley trace, which allow both fixed and Gaussian predictors in the same model. Section 3 describes the new unconditional power approximation for designs with one or more Gaussian covariates. Section 4 compares empirical results with the new method and existing approaches. Section 5 details the applied example in oral cancer. Section 6 provides concluding remarks.

2. Notation, Models, and Hypothesis Testing

2.1. Notation

Throughout, we follow the notation of Muller and Stewart (2006). Define $\mathbf{a}$ to be an $(n \times 1)$ column vector. Let $\mathbf{A} = \{a_{ij}\}$ be an $(n \times m)$ matrix with transpose $\mathbf{A}' = \{a_{ji}\}$. Define $\mathbf{A} \otimes \mathbf{B}$ to be the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. Let $\mathbf{I}_p$ be a $(p \times p)$ identity matrix. Define $\mathbf{1}_n$ to be an $(n \times 1)$ vector with each element equal to 1. Let $\mathbf{L}(r, c) = \{l_{ij}\}$ be an $(r \times c)$ matrix with $l_{1,1} = 1$ and $l_{i,j} = 0$ for all $(i, j) \neq (1, 1)$. Let $E(\mathbf{A})$ be the expectation of the matrix $\mathbf{A}$. Write the covariance of the vector $\mathbf{a}$ as $\mathcal{V}(\mathbf{a})$.

Let $\mathcal{F}(f; \nu_n, \nu_d, \omega)$ be the cumulative distribution function of a noncentral $\mathcal{F}$ distribution with ν_n numerator degrees of freedom, ν_d denominator degrees of freedom, and noncentrality parameter ω . Define $\mathcal{F}^{-1}$ such that for $0 < p < 1$

$$\mathcal{F}(f; \nu_n, \nu_d, \omega) = p \Leftrightarrow \mathcal{F}^{-1}(p; \nu_n, \nu_d, \omega) = f. \quad (1)$$

For $\boldsymbol{\mu}$ an $(p \times 1)$ vector, $\boldsymbol{\Sigma}$ a $(p \times p)$ symmetric, positive definite matrix, and $\mathbf{a}$ a $(p \times 1)$ random vector, let $\mathbf{a} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ indicate that $\mathbf{a}$ has a p -dimensional vector Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ (Arnold, 1981). For $\mathbf{M}$ an $(N \times p)$ matrix, $\boldsymbol{\Xi}$ and $\boldsymbol{\Sigma}$ symmetric, positive definite matrices of dimension $(N \times N)$ and $(p \times p)$, respectively, and $\mathbf{A}$ an $(N \times p)$ random matrix, write $\mathbf{A} \sim \mathcal{N}_{N,p}(\mathbf{M}, \boldsymbol{\Xi}, \boldsymbol{\Sigma})$ to indicate that $\mathbf{A}$ has a matrix Gaussian distribution with mean $\mathbf{M}$, column covariance $\boldsymbol{\Xi}$, and row covariance $\boldsymbol{\Sigma}$ (Arnold, 1981).

Let

$$CS_p(\rho) = [\mathbf{1}_p \mathbf{1}_p' \rho + \mathbf{I}_p (1 - \rho)] \quad (2)$$

be a $(p \times p)$ compound symmetric correlation matrix.

For $0 < \rho < 1$ and $0 < \delta$, define

$$LEAR_p(\rho, \delta) = \rho^{d_{\min} + \delta[(d_{jk} - d_{\min})/(d_{\max} - d_{\min})]} \quad (3)$$

to be a $(p \times p)$ LEAR correlation structure, with d_{jk} the distance between the j^{th} and k^{th} measurement, $d_{\min} = \min_{\{j < k\}} d_{jk}$ and $d_{\max} = \max_{\{j < k\}} d_{jk}$ (Simpson et al., 2010).

2.2. The General Linear Multivariate Model with Gaussian Covariates

Let $\mathbf{X}$ be the $(N \times q)$ matrix of predictor variables. Let $\mathbf{Y}$ be the $(N \times p)$ matrix of outcomes, such that for all $i \in \{1, \dots, N\}$, $[\text{row}_i(\mathbf{Y})]' \sim \mathcal{N}_p(\mathbf{X}\mathbf{B}, \mathbf{\Sigma}_Y)$. Define $\mathbf{B}$ to be the $(q \times p)$ matrix of regression coefficients. Let $\mathbf{E}$ be the $(N \times p)$ matrix of errors such that for all $i \in \{1, \dots, N\}$, $[\text{row}_i(\mathbf{E})]' \sim \mathcal{N}_p(0, \mathbf{\Sigma}_E)$.

The general linear multivariate model (GLMM) is

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}. \quad (4)$$

For GLMMs with fixed and Gaussian predictors, partition $\mathbf{X}$ and $\mathbf{B}$ into fixed and random submatrices. Define $\mathbf{F}$ to be the $(N \times q_F)$ matrix of fixed predictors. Let $\mathbf{G}$ be the $(N \times q_G)$ matrix of random predictors, such that for all $i \in \{1, \dots, N\}$, $[\text{row}_i(\mathbf{G})]' \sim \mathcal{N}_{q_G}(0, \mathbf{\Sigma}_G)$. Then $\mathbf{X} = [\mathbf{F} \ \mathbf{G}]$. Without loss of generality, assume that $\mathbf{X}$ is full rank with $r = \text{rank}(\mathbf{X}) = (q_F + q_G)$. Note that r depends on the number of random Gaussian covariates. Define $\mathbf{\Sigma}_{YG}$ to be the $(p \times q_G)$ matrix such that

$$\mathcal{V}(\{\text{row}_i[\mathbf{Y} \ \mathbf{G}]\}) = \begin{bmatrix} \mathbf{\Sigma}_Y & \mathbf{\Sigma}_{YG} \\ \mathbf{\Sigma}_{GY} & \mathbf{\Sigma}_G \end{bmatrix}. \quad (5)$$

Let $\mathbf{B}_F$ be the $(q_F \times p)$ matrix of regression coefficients associated with fixed predictors, and let $\mathbf{B}_G$ be the $(q_G \times p)$ matrix of regression coefficients associated with random predictors. Then $\mathbf{B} = [\mathbf{B}_F' \ \mathbf{B}_G']'$.

The GLMM with Gaussian predictors [GLMM(F, G)] is

$$\mathbf{Y} = \mathbf{F}\mathbf{B}_F + \mathbf{G}\mathbf{B}_G + \mathbf{E}. \quad (6)$$

2.3. The General Linear Hypothesis

We define the general linear hypothesis for the GLMM(F, G). Define $\mathbf{C}$ to be the $(a \times q)$ matrix of between participant contrasts. Let $\mathbf{U}$ be the $(p \times b)$ matrix of within participant contrasts. Define the $(a \times b)$ matrix $\mathbf{\Theta} = \mathbf{C}\mathbf{B}\mathbf{U}$, and let

$$s = \min(a, b). \quad (7)$$

The null and alternative hypotheses can be stated as

$$H_0: \boldsymbol{\Theta} = \boldsymbol{\Theta}_0 \quad (8)$$

$$H_A: \boldsymbol{\Theta} \neq \boldsymbol{\Theta}_0. \quad (9)$$

Partition $\mathbf{C}$ into submatrices corresponding to $\mathbf{B}_F$ and $\mathbf{B}_G$. Define $\mathbf{C}_F$ to be the $(a \times q_F)$ submatrix of contrast coefficients associated with $\mathbf{B}_F$. Define $\mathbf{C}_G$ to be the $(a \times q_G)$ submatrix of contrast coefficients associated with $\mathbf{B}_G$. Then $\mathbf{C} = [\mathbf{C}_F \ \mathbf{C}_G]$. Throughout this manuscript, we assume that the hypothesis of interest deals with fixed predictors, implying that $\mathbf{C}_G \equiv \mathbf{0}$.

Several test statistics have been described for the general linear hypothesis (Muller and Stewart, 2006). We restrict our attention to the Hotelling-Lawley trace, given its correspondence to the Wald test in balanced mixed models (Muller et al., 2007). Define $\nu_e = N - r$. The Hotelling-Lawley trace is

$$t_{HLT} = \text{tr}(\mathbf{S}_h \mathbf{S}_e^{-1}) \quad (10)$$

where

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}, \quad (11)$$

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\mathbf{B}}, \quad (12)$$

$$\hat{\boldsymbol{\Sigma}} = (\mathbf{Y} - \hat{\mathbf{Y}})'(\mathbf{Y} - \hat{\mathbf{Y}})/\nu_e, \quad (13)$$

$$\mathbf{M} = \mathbf{C}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{C}', \quad (14)$$

$$S_h = (\widehat{\Theta} - \Theta_0)' M^{-1} (\widehat{\Theta} - \Theta_0), \quad (15)$$

and

$$S_e = \nu_e U' \widehat{\Sigma} U. \quad (16)$$

We note that $S_h S_e^{-1}$ will have a single non-zero eigenvalue for $s = 1$.

3. Approximate Unconditional Power for the General Linear Multivariate Model with One or More Gaussian Predictors

Our unconditional power approximation extends the noncentral $\mathcal{F}$ power approximation described by Muller *et al.* (1992) for the Hotelling-Lawley trace. As in Muller *et al.* (1992), specify the Type I error rate, α , and the design, including F , B_F , C_F , U , Θ_0 , Σ_G , Σ_{YG} , and Σ_Y . As in Sampson (1974), calculate the error covariance

$$\Sigma_E = \Sigma_Y - \Sigma_{YG} \Sigma_G^{-1} \Sigma_{GY}. \quad (17)$$

Compute the corresponding population error sum of squares as

$$S_e = \nu_e U' \Sigma_E U.$$

For $C_G = \mathbf{0}$, $\Theta = C_F B_F U$.

Define

$$g(\nu_e, a, b) = \frac{[\nu_e^2 - \nu_e(2b + 3) + b(b + 3)](ab + 2)}{\nu_e(a + b + 1) - (a + 2b + b^2 - 1)} + 4. \quad (18)$$

Note that in Equation 18, ν_e is a function of q_G , the number of random covariates. An equation similar to Equation 18 appears in Muller *et al.*, (2007). In Muller *et al.*, (2007), ν_e depends only on the rank of X , which is assumed to be fixed, making $q_G = 0$.

Calculate approximate unconditional power as follows. Find the critical value under the null using

$$F_{crit} = \mathcal{F}_F^{-1}[1 - \alpha; ab, g(\nu_e, a, b), 0]. \quad (19)$$

We now calculate the scaled population hypothesis sum of squares. The scaling arises from a Taylor series approximation based on the matrix-variate beta distribution of type I. The derivation of the Taylor series is given in Appendix A. Calculate

$$S_{h*} = \left(\frac{N - q_G}{N} \right) (\Theta - \Theta_0)' [C_F (F' F)^{-1} C_F']^{-1} (\Theta - \Theta_0). \quad (20)$$

Compute the noncentrality parameter as

$$\omega = \nu_e \text{tr}(S_{h*} S_e^{-1}). \quad (21)$$

Finally, calculate power as

$$\text{Power} \approx 1 - \mathcal{F}_F[F_{crit}; ab, g(\nu_e, a, b), \omega]. \quad (22)$$

Note that Equation 22 appears in Muller *et al.* (1992), with the exception of alternative formulae for the denominator degrees of freedom and the non-centrality parameter.

The power calculation is implemented in a free, open-source rPowerlib package published at <http://github.com/SampleSizeShop/rPowerlib>.

4. Simulation Experiment

4.1. Methods

We compared the new approximation described in Section 3, several existing approaches (Glueck and Muller, 2003; Shieh, 2005; Muller *et al.*, 1992), and empirical unconditional power.

Note that only the method of Shieh (2005) allows for multiple Gaussian predictors. To implement the approach of Glueck and Muller (2003), we calculated power as if the model had contained only a single covariate. From the full set of covariates, we selected the single covariate which had the strongest correlation with any of the outcomes. For the approach described by Muller *et al.* (2007), we calculated power as if the design had only contained fixed predictors. We assumed that $\Sigma_E = \Sigma_Y$, $X = F$, $B = B_F$ and $C = C_F$.

We calculated empirical power using a two-level Monte Carlo simulation. For each study design, we fixed $\alpha, F, B_F, C_F, U, \Theta_0, \Sigma_G, \Sigma_{YG}$, and Σ_Y . We calculated Σ_E as in Equation 17. For

the first level, we generated 1,000 realizations of $\mathbf{G}$. In the second level, for each realization of $\mathbf{G}$, we simulated 1,000 replicates of $\mathbf{E}$. For each realization of $\mathbf{G}$, we formed $\mathbf{X} = [\mathbf{F} \ \mathbf{G}]$. For each $\mathbf{G}$, associated $\mathbf{X} = [\mathbf{F} \ \mathbf{G}]$ and $\mathbf{E}$, we calculated $\mathbf{Y} = \mathbf{F}\mathbf{B}_F + \mathbf{G}\mathbf{B}_G + \mathbf{E}$ as in Equation 6. Empirical power was calculated as the proportion of replicates for which the null hypothesis was rejected. The process produced a list of 1000 empirical power values, one for each $\mathbf{G}$. We estimated empirical unconditional power as the mean of the power values obtained for each realization of $\mathbf{G}$.

A timing experiment was conducted to compare total CPU time for empirical unconditional power and for the novel approximation. Timing results were obtained on a Macbook Pro running OSX Mavericks. The machine had 16GB of RAM and a 2.5 GHz Dual-core Intel Core i5 processor. We used the R (R Core Team, 2014) function `proc.time()` to calculate CPU time to the nearest millisecond for the approximate and empirical power calculations.

We used R version 3.0.2 (R Core Team, 2014) to calculate empirical power and to compute power using the new approximation (Section 3) and the method of Shieh. The JavaStatistics library from the GLIMMPSE software product (Kreidler et al., 2013) was used to calculate power for fixed designs (Muller et al., 1992) and single covariate designs (Glueck and Muller, 2003). Source code and instructions for reproducing the manuscript results are available at <http://github.com/SampleSizeShop/rPowerlib>.

Study Designs—We validated the methods for several study designs and hypotheses. In each design, treatment was the primary fixed predictor, with either two or four treatment groups. We included designs with either one, five, or six outcomes. We considered designs and hypotheses for both $s = 1$ and $s > 1$ (see Equation 7). Design matrices, contrasts, and covariance matrices used in the experiment are shown in Appendix B. For each design, we varied the following parameters: 1) the number of covariates, $q_G \in \{1, 3, 6\}$, 2) the per treatment sample size, $N_t \in \{10, 100\}$, and 3) a scale factor applied to Σ_{YG} , $k \in \{0.5, 1, 1.5, 2\}$. As k increases, the residual error is more strongly reduced by the covariates. In all cases, $\mathbf{B}_F = d \times \mathbf{L}(q, p)$, with d selected such that the power calculated using the approximation described in Section 3 was 0.9. Values used for Σ_G and Σ_{YG} are shown in Appendix B.

Performance Criteria—For each design and power method, we computed the deviation as approximate power minus empirical power. We produced box plots summarizing the deviations across all designs in the validation study and stratified by 1) sample size, 2) the number of random covariates, and 3) k , the scale factor for Σ_{YG} .

Positive deviations indicated that the approximate power values were larger than the empirical power values. Negative deviations indicated that the approximate power values were smaller than the empirical power values.

4.2. Results

In general, the new approximation provided the most accurate power values (Figure 1). All four methods had average deviations close to zero. The method described by Shieh (2005) occasionally overestimated power by up to 10 percentage points. The methods which ignored covariates or retained only a single covariate tended to produce power values which

were too low. The maximum absolute deviations observed were 0.02 for the new method, 0.10 for the method of Shieh, 0.46 for the single covariate approach, and 0.59 for the fixed only method.

The new method remained accurate regardless of the number of random covariates (Figure 2). The Shieh method lost accuracy as the number of covariates increased. The single covariate and fixed only methods showed the greatest degradation as the number of covariates increased.

With increasing sample size (Figure 3), the deviations from empirical power improved for all methods, although the single covariate and fixed only approaches continued to have extreme deviations for some study designs. The new approximation was accurate for both small ($N = 10$) and large ($N = 100$) samples, with increasing accuracy as the sample size increased. The Shieh method improved with increasing sample size, a reasonable result given its dependence on asymptotic assumptions.

The new method remained accurate as the correlation between the covariates and the outcomes increased (Figure 4). The method of Shieh overestimated power, but the overestimation remained constant as the correlation increased. Although the average deviations for the single covariate and the fixed methods remained close to zero, extreme outliers indicated that the methods were less reliable as the covariates became more important in the model.

Simulation times ranged from 29 to 588 seconds, depending on the complexity of the design. The maximum calculation time was 2 milliseconds, with most times falling below the single millisecond resolution of the `proc.time()` function in R (R Core Team, 2014). Mean simulation and calculation times for designs with the same number of covariates, per group sample size, and repeated measures are summarized in Table 1.

5. Example Power Calculation for a Hypothetical Study of Salivary Biomarkers in Oral Cancer

We demonstrate a power calculation for a hypothetical experiment investigating salivary biomarkers in oral cancer. Elashoff *et al.* (2012) showed that IL-8 and other mRNA markers were elevated in individuals with oral squamous cell carcinoma. A natural extension to the Elashoff *et al.* (2012) study would be a clinical trial to determine if treatment reduces the expression of these salivary biomarkers.

We present a power analysis for this hypothetical study. Participants are randomized to one of two treatments, and salivary biomarkers are measured at baseline, six months, and 1 year. The Gaussian covariates are age and baseline IL-8 level. The hypothesis of interest is the test of the time by treatment interaction. We use the Hotelling-Lawley trace at $\alpha = 0.05$.

With 35 participants in each treatment group, the fixed portion of the design matrix is

$$\mathbf{F} = \mathbf{1}_{35} \otimes \mathbf{I}_2. \quad (23)$$

We obtained reasonable values for $\mathbf{B}_F$ from Table 3 in Elashoff *et al.* (2012). The average difference in IL-8 between cases and controls was 2.14. Under the alternative hypothesis, we assume that the difference will be near zero by the one year measurement for the treatment group, but that the difference will remain constant across time for the control group. The fixed portion of the $\mathbf{B}$ matrix is

$$\mathbf{B}_F = \begin{bmatrix} 2.14 & 0 \\ 2.14 & 2.14 \end{bmatrix}. \quad (24)$$

We define the between-participant contrast as

$$\mathbf{C}_F = [1 \ -1], \quad (25)$$

and the within-participant contrast as

$$\mathbf{U} = [1 \ -1]'. \quad (26)$$

We assume a null hypothesis of no interaction,

$$\boldsymbol{\Theta}_0 = [0]. \quad (27)$$

We obtained covariance values from Tables 1 and 3 in Elashoff *et al.* (2012). Ignoring covariates, we assumed that the standard deviation of IL-8 was 2.2 (variance 4.84). We have no information regarding the correlation of IL-8 levels over time. We anticipated that the correlation between the 6 month measurement and the 1 year measurement was about 0.4. Thus

$$\boldsymbol{\Sigma}_Y = 4.84 \times \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix}. \quad (28)$$

For the covariates, we assumed that the standard deviation for age was 12.7 years (variance 161.29 years), and 2.2 (variance 4.84) for IL-8. We assumed that age and the IL-8 level were independent, yielding the following $\boldsymbol{\Sigma}_G$ matrix:

$$\Sigma_G = \begin{bmatrix} 4.84 & 0 \\ 0 & 161.29 \end{bmatrix}. \quad (29)$$

We assumed that correlation between the baseline IL-8 and the 6 month measure was 0.4 and that the correlation between the baseline IL-8 measure and the 1-year measure was 0.2. The correlation between age and IL-8 was assumed to be 0.1 for both the 6 month and 1 year follow-up. After converting the correlation values to covariances, we obtained

$$\Sigma_{YG} = \begin{bmatrix} 0.4(4.84) & 0.1\sqrt{4.84(161.29)} \\ 0.2(4.84) & 0.1\sqrt{4.84(161.29)} \end{bmatrix} \quad (30)$$

$$= \begin{bmatrix} 1.94 & 2.79 \\ 0.97 & 2.79 \end{bmatrix}. \quad (31)$$

Then, using Equation 17, the reduced covariance is

$$\Sigma_E = \begin{bmatrix} 4.01 & 1.50 \\ 1.50 & 4.60 \end{bmatrix}. \quad (32)$$

The matrices defined above were input into the rPowerlib package at <http://github.com/SampleSizeShop/rPowerlib>. The software calculated power using the method described in Section 3. At an α level of 0.05, power was computed to be 0.956 for the Hotelling-Lawley trace test of time by treatment interaction, with an interaction magnitude of 2.14.

6. Discussion

We present a power approximation for the general linear multivariate model with one or more Gaussian covariates. The results apply to both single and multiple parameter hypotheses. The method has excellent accuracy, both for small and large samples. In addition, the method performs better than previous asymptotic results described by Shieh. Both the new method and the method of Shieh are more accurate than methods which failed to account for multiple covariates.

The timing experiment shown in Table 1 demonstrates that the new method is several orders of magnitude faster than a simulation based approach for a single calculation. Note that in a typical power and sample size calculation, one would calculate not a single point, but multiple points, comprising a smooth power curve. This further reduces the impetus to conduct a simulation. However, we acknowledge that some users may still require simulation-based power results, and so provide both empirical power via simulation and fast approximate power in the R package associated with this manuscript.

Researchers faced with the task of designing studies with multiple Gaussian covariates may find the collection of the required inputs daunting. However, as shown in Section 5, it is possible to obtain reasonable values for mean differences and variability fairly easily from the literature. Standard deviations are commonly reported and can be used to calculate the required covariance matrices. Correlations can be more difficult to find, but can often be obtained from pilot data. When neither of these options is available, clinical expertise can be used to fill in the gaps.

The power approximation has several limitations. The results assume complete and balanced data. In addition, the method may not be applicable to designs in which treatment assignment depends on the values of covariates. Finally, the power approximation was derived for tests of only fixed predictors. We plan further work to extend the analytic results to hypotheses involving both fixed and random predictors.

Future research is needed to determine whether a similar approximation would retain accuracy for non-Gaussian random predictors with finite second moments. The proposed approach relies on a Taylor series approximation using the matrix-variate beta distribution of type I. The matrix-variate beta distribution of type I arises only because of the assumption that the covariates follow a matrix Gaussian distribution. Should the random covariates instead follow a non-Gaussian distribution with finite second moments, it is far from clear what an appropriate approximating distribution for the non-centrality might be. Detailed analytic work and confirmatory simulation studies for non-Gaussian predictors are necessary.

Acknowledgments

Funding for this work was provided by two grants: NIDCR 1 R01 DE020832-01A1 and NIGMS 9R01GM121081-05. The NIDCR grant was awarded to the University of Florida (Keith E. Muller, PI; Deborah H. Glueck, University of Colorado site PI). The NIGMS grant was awarded to the Colorado School of Public Health, D. H. Glueck, Principal Investigator. The content of this paper is solely the responsibility of the authors, and does not necessarily represent the official views of the National Institute of Dental and Craniofacial Research, National Institute of General Medical Science nor National Institutes of Health. A portion of this paper was submitted to the University of Colorado Denver in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biostatistics for Dr. Sarah M. Kreidler.

Bibliography

- Arnold, SF. The Theory of Linear Models and Multivariate Analysis. Wiley; New York: 1981.
- Elashoff D, Zhou H, Reiss J, Wang J, Xiao H, Henson B, Hu S, Arellano M, Sinha U, Le A, Messadi D, Wang M, Nabili V, Lingen M, Morris D, Randolph T, Feng Z, Akin D, Kastratovic DA, Chia D, Abemayor E, Wong DTW. 2012; Prevalidation of salivary biomarkers for oral cancer detection. *Cancer Epidemiology, Biomarkers & Prevention*. 21(4):664–672.
- Glueck DH, Muller KE. 2003; Adjusting power for a baseline covariate in linear models. *Statistics in Medicine*. 22(16):2535–2551. [PubMed: 12898543]
- Gupta, AK, Nagar, DK. Matrix Variate Distributions. Chapman & Hall; Boca Raton, FL: 2000. Appendix A
- Kreidler SM, Muller KE, Grunwald G, Ringham BM, Coker-Dukowitz ZT, Sakhadeo UR, Barón AE, Glueck DH. 2013; GLIMPSE: Online power computation for linear models with and without a baseline covariate. *Journal of Statistical Software*. 54(10)
- Muller KE, Barton CN. 1989; Approximate power for repeated-measures ANOVA lacking sphericity. *Journal of the American Statistical Association*. 84(406):549–555.

- Muller KE, Benignus VA. 1992; Increasing scientific power with statistical power. *Neurotoxicology and Teratology*. 14(3):211–219. [PubMed: 1386138]
- Muller KE, Edwards LJ, Simpson SL, Taylor DJ. 2007; Statistical tests with accurate size and power for balanced linear mixed models. *Statistics in Medicine*. 26(19):3639–3660. [PubMed: 17394132]
- Muller KE, Lavange LM, Ramey SL, Ramey CT. 1992; Power calculations for general linear multivariate models including repeated measures applications. *Journal of the American Statistical Association*. 87(420):1209–1226. [PubMed: 24790282]
- Muller KE, Peterson BL. 1984; Practical methods for computing power in testing the multivariate general linear hypothesis. *Computational Statistics and Data Analysis*. 2:143–158.
- Muller, KE, Stewart, PW. *Linear Model Theory: Univariate, Multivariate, and Mixed Models*. John Wiley and Sons; Hoboken, New Jersey: 2006.
- R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2014.
- Sampson AR. 1974; A tale of two regressions. *Journal of the American Statistical Association*. 69(347):682–689.
- Shieh G. 2005; Power and sample size calculations for multivariate linear models with random explanatory variables. *Psychometrika*. 70(2):347–358.
- Simpson SL, Edwards LJ, Muller KE, Sen PK, Styner MA. 2010; A linear exponent AR(1) family of correlation structures. *Statistics in Medicine*. 29(17):1825–1838. [PubMed: 20658550]

Appendix A. Scale Factor for the Hypothesis Sum of Squares

We use a first order Taylor series approximation to obtain the scale factor for the hypothesis sum of squares. Under the assumptions of Section 2, and with $\mathbf{H}_G = \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'$,

$$\mathbf{M} = \mathbf{C}_F [\mathbf{F}' (\mathbf{I}_N - \mathbf{H}_G) \mathbf{F}]^{-1} \mathbf{C}_F'. \quad (\text{A.1})$$

From Gupta and Nagar (2000, p. 168, Theorem 5.2.4), $\mathbf{H}_G$ is distributed $B_N^I \{(q_G/2), [(N - q_G)/2]\}$, with B_N^I the matrix-variate beta distribution of type I. The expectation of $\mathbf{H}_G$ is (Gupta and Nagar, 2000, p. 178, Theorem 5.3.19)

$$E(\mathbf{H}_G) = \frac{q_G}{N} \mathbf{I}_N. \quad (\text{A.2})$$

Replace $\mathbf{H}_G$ with $E(\mathbf{H}_G)$ in Equation A.1 to obtain

$$\mathbf{M} \approx \frac{N}{N - q_G} \mathbf{C}_F (\mathbf{F}' \mathbf{F})^{-1} \mathbf{C}_F'. \quad (\text{A.3})$$

Then the result follows by substitution into Equation 15.

Appendix B. Design Matrices for the Simulation Experiment

We describe the inputs for the simulation experiment in Section 4. The matrix representation of each study design is shown in Table B.2.

The dimensions of Σ_G and Σ_{YG} were determined by the number of covariates and the dimensions of Σ_Y in a given design. In addition, the values for Σ_G and Σ_{YG} were selected to ensure that the resulting Σ_e matrix was positive definite (Equation 17), given the structure of Σ_Y . Values for Σ_G are shown in Table B.3. We summarize the values of Σ_{YG} in Table B.4, with

$$\begin{aligned} \mathbf{a} &= [1 \ 0.75 \ 0.5]', \\ \mathbf{b} &= [1 \ 0.9 \ 0.8 \ 0.7 \ 0.6 \ 0.5]', \\ \mathbf{c} &= [0.542 \ 0.477 \ 0.420 \ 0.369 \ 0.325]' \end{aligned}$$

Table B.2

Matrices for designs, regression coefficients, contrasts, and covariances

X_F	C	B_F	U	Σ_Y
$I_2 \otimes \mathbf{1}_{N_t}$	$[1 \ -1]$	$d \times L(2, 1)$	$[1]$	$\sigma_Y^2 = 5$
$I_2 \otimes \mathbf{1}_{N_t}$	$[1 \ -1]$	$d \times L(2, 5)$	$\begin{bmatrix} -0.63 & 0.53 & -0.32 & 0.12 \\ -0.32 & -0.27 & 0.63 & -0.48 \\ 0.00 & -0.53 & 0.00 & 0.72 \\ 0.32 & -0.27 & -0.63 & -0.48 \\ 0.63 & 0.53 & 0.32 & 0.12 \end{bmatrix}$	$5 \times CS_5 (0.04)$
$I_2 \otimes \mathbf{1}_{N_t}$	$[1 \ -1]$	$d \times L(2, 5)$	$\begin{bmatrix} -0.63 & 0.53 & -0.32 & 0.12 \\ -0.32 & -0.27 & 0.63 & -0.48 \\ 0.00 & -0.53 & 0.00 & 0.72 \\ 0.32 & -0.27 & -0.63 & -0.48 \\ 0.63 & 0.53 & 0.32 & 0.12 \end{bmatrix}$	$5 \times LEAR_5 (0.4, 1)$
$I_2 \otimes \mathbf{1}_{N_t}$	$[1 \ -1]$	$d \times L(2, 6)$	$\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$5 \times [CS_3 (0.1) \otimes CS_2 (0.2)]$
$I_4 \otimes \mathbf{1}_{N_t}$	$[\mathbf{1}_3 \ -I_3]$	$d \times L(4, 1)$	$[1]$	$\sigma_Y^2 = 5$
$I_4 \otimes \mathbf{1}_{N_t}$	$[\mathbf{1}_3 \ -I_3]$	$d \times L(4, 5)$	$\begin{bmatrix} -0.63 & 0.53 & -0.32 & 0.12 \\ -0.32 & -0.27 & 0.63 & -0.48 \\ 0.00 & -0.53 & 0.00 & 0.72 \\ 0.32 & -0.27 & -0.63 & -0.48 \\ 0.63 & 0.53 & 0.32 & 0.12 \end{bmatrix}$	$5 \times CS_5 (0.04)$

X_F	C	B_F	U	Σ_Y
$I_4 \otimes \mathbf{1}_{N_t}$	$[\mathbf{1}_3 \ -I_3]$	$d \times L(4, 5)$	$\begin{bmatrix} -0.63 & 0.53 & -0.32 & 0.12 \\ -0.32 & -0.27 & 0.63 & -0.48 \\ 0.00 & -0.53 & 0.00 & 0.72 \\ 0.32 & -0.27 & -0.63 & -0.48 \\ 0.63 & 0.53 & 0.32 & 0.12 \end{bmatrix}$	$5 \times LEAR_5(0.4, 1)$
$I_4 \otimes \mathbf{1}_{N_t}$	$[\mathbf{1}_3 \ -I_3]$	$d \times L(4, 6)$	$\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$5 \times [CS_3(0.1) \otimes CS_2(0.2)]$

and

$$d = [0.36 \ 0.30 \ 0.36 \ 0.30 \ 0.36 \ 0.30]'$$

Table B.3

Values for Σ_G by number of covariates

Number of Covariates	Σ_G
1	$\sigma_G^2 = 1$
3	$\begin{bmatrix} 0.91 & 0.01 & 0.01 \\ 0.01 & 1.21 & 0.01 \\ 0.01 & 0.01 & 1.51 \end{bmatrix}$
6	$\begin{bmatrix} 0.91 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 1.21 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 1.51 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 1.81 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 2.11 & 0.01 \\ 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 2.41 \end{bmatrix}$

Table B.4

Values for Σ_{YG} for a given Σ_Y and number of covariates

Σ_Y	Σ_{YG}		
	1 covariate	3 covariates	6 covariates
$\sigma_Y^2 = 5$	$\sigma_Y^2 = 0.6$	$0.6 \otimes a'$	$0.6 \otimes b'$
$5 \times CS_5(0.04)$	$0.2 \otimes \mathbf{1}_5$	$0.2 \otimes \mathbf{1}_5 \otimes a'$	$0.2 \otimes \mathbf{1}_5 \otimes b'$
$5 \times LEAR_5(0.4, 1)$	c	$c \otimes a'$	$c \otimes b'$
$5 \times [CS_3(0.1) \otimes CS_2(0.2)]$	d	$d \otimes a'$	$d \otimes b'$

Figure 1.
Deviation From Empirical Power Overall

Figure 2.
Deviation From Empirical Power by the Number of Random Covariates

Figure 3.
Deviation From Empirical Power by Per Group Sample Size

Figure 4.
Deviation From Empirical Power by Σ_{YG} Scale Factor

Table 1

Comparison of simulation time to calculation time

Number of covariates	Per group sample size	Number of repeated measures	Mean simulation time (sec.)	Mean calculation time (sec. $\times 10^{-4}$)	Number of designs considered
1	10	1	35	3	8
1	10	5	55	3	16
1	10	6	58	2	8
1	100	1	159	2	8
1	100	5	324	1	16
1	100	6	367	3	8
3	10	1	37	1	8
3	10	5	59	2	16
3	10	6	63	3	8
3	100	1	173	4	8
3	100	5	347	3	16
3	100	6	400	4	8
6	10	1	41	4	8
9	10	5	64	3	16
9	10	6	69	4	8
9	100	1	202	1	8
9	100	5	389	4	16
9	100	6	441	3	8