Approaches to Estimate Between- and Within-Subject Correlation Coefficients in Longitudinal Repeated-Measures Studies

Jennifer N. Cooper[1], Jason P. Sulkowski[1,2], Katherine J. Deans[1,2]
Peter C. Minneci[1,2]
[1]Center for Surgical Outcomes Research and Center for Innovation in Pediatric Practice, The Research Institute at Nationwide Children's Hospital, 700 Childrens Dr., Columbus, OH 43205
[2]Department of Surgery, Nationwide Childrens Hospital, 700 Childrens Dr., Columbus, OH 43205

Abstract
Multivariate repeated-measures data offer a unique opportunity to examine the joint progression of multiple variables over time. Several distinct types of correlations between variables can be derived from such studies. The overall correlation between two variables can be decomposed into within- and between-subject correlations that reflect associations at the individual and collective levels. A number of different models are useful for estimating these correlations, with bivariate and multivariate linear mixed models being most commonly used. Currently, the SAS MIXED procedure has few options for Kronecker product covariance structures and all best accommodate equally spaced measurements. Univariate mixed models with separate terms for between- and within-subject associations are an alternative that allows for unequally spaced measurements. We explore differences in overall, within- and between-subject correlation coefficients derived from multivariate, bivariate, and univariate mixed models under a variety of covariance structures. We illustrate these differences using data on markers of intravascular hemolysis and nitric oxide consumption in children treated with extracorporeal membrane oxygenation.

Key Words: mixed model; longitudinal; covariance structure; pediatrics

1. Introduction

In most longitudinal studies, subjects have measurements on more than one response over time. With such multivariate longitudinal data, there are a number of research questions that can be asked regarding associations between the responses. Importantly, however, any method used to answer these questions should account for the dependence of the observations. Namely, it is expected that repeated measurements on any single response will be correlated over time while multiple responses on the same subject will be correlated at any given time point.

Bland and Altman were the first to consider the estimation of the correlation between two continuous variables with repeated measurements, and they introduced the idea that the correlation coefficient between two such variables can be considered to have two components: between-subject correlation and within-subject correlation (1, 2). The

between-subject correlation indicates the correlation between subject means. It measures whether subjects with high values of one variable also generally have high values of the other variable over the course of the study. The within-subject correlation describes whether an increase in one variable over time within a subject is associated with an increase in the other variable. Several researchers have recently investigated this problem through the use of maximum likelihood estimation using specialized software (3) or bivariate linear mixed models (LMM) fit using SAS PROC MIXED (SAS Institute Inc., Cary, NC) (4-6).

In addition to within- and between-subject correlation coefficients, other types of correlation coefficients can be estimated from bivariate linear mixed models. Fieuws et al. described two other types of correlations as 1) how the evolution of one response is related to the evolution of another response ('association of the evolutions') and how the association between responses evolves over time ('evolution of the association') (7, 8). These types of correlations can be estimated from bivariate linear mixed models as well, but they seem to be most interpretable if the responses display monotonic, approximately linear, changes over time; otherwise it may become difficult to interpret the correlation between evolutions in responses over time. Fieuws et al. assumed conditional independence of the observations given the random effects, but this is not a necessary restriction, and is in fact unreasonable for some datasets (7, 8).

Other types of joint modelling approaches exist for bivariate longitudinal data, such as treating one response as a time-varying covariate while modelling the other as the dependent variable. This approach is used extensively in the literature, and is particularly common for jointly analyzing time-dependent responses and a survival outcome (9). This approach does address the correlation between responses in the sense that the direction of the corresponding regression coefficient mirrors the direction of the correlation between the two variables, and standard significance tests for the coefficients can be used to assess the strength of the association. However, it is problematic to use responses as predictors because the choice of the response to treat as the dependent variable is often arbitrary, and different choices may in fact lead to different conclusions (10). Also, additional computations are required to determine marginal responses and marginal associations after fitting such models (7). Another possible strategy to analyze bivariate longitudinal data is to fit a marginal model for each response, then join these models in some way (11, 12). However, with this approach, it is not possible to estimate correlations between responses because this strategy does not explicitly model the correlation structure of the data (10).

All of the above concepts can be easily extended to the analysis of multivariate rather than bivariate longitudinal data (10). Because there is no consensus on the best strategy for the analysis of multivariate longitudinal data, and the chosen strategy to estimate correlations between multiple longitudinal responses will naturally depend on the research question, the aim of this study was to estimate and explore differences in overall, within- and between-subject correlations, as well as correlations representing the 'association of the evolutions' from multivariate, bivariate, and univariate linear mixed models under a variety of variance covariance patterns.

2. Methods

2.1 Bivariate and multivariate linear mixed modelling strategies

For a bivariate (q=2 responses) or multivariate (q>2 responses) LMM, with a maximum of p successive time points for each subject, we represent each subject's responses $\boldsymbol{y}$ by a vector of length qp x 1 by stacking first all responses at each time point, then stacking the consecutive time points. It is assumed that $\boldsymbol{y}$ follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and qp x qp positive definite variance covariance matrix $\boldsymbol{\Omega}$. For an individual subject, denoted subject i of N total subjects, with measurements at m_i number of time points, where $m_i<p$, the dimensions of the response vector and variance covariance matrix will be qm_i and qm_i x qm_i respectively.

The general form of a LMM is:

$$\boldsymbol{Y_i} = \boldsymbol{X_i\beta} + \boldsymbol{Z_ib_i} + \boldsymbol{\varepsilon_i},$$
$$\boldsymbol{b_i} \sim N(\boldsymbol{0, D}),$$
$$\boldsymbol{\varepsilon_i} \sim N(\boldsymbol{0, R_i})$$

where $\boldsymbol{b_1}, \boldsymbol{b_2},\ldots, \boldsymbol{b_N}, \boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_2},\ldots,\boldsymbol{\varepsilon_N}$ are independent. $\boldsymbol{X_i}$ and $\boldsymbol{Z_i}$ are design matrices of known covariates, $\boldsymbol{\beta}$ is a vector containing the fixed effects, $\boldsymbol{b_i}$ is a vector containing the random effects, $\boldsymbol{\varepsilon_i}$ is a vector of residual components, $\boldsymbol{D}$ is the variance covariance matrix of the random effects, and $\boldsymbol{R_i}$ the variance covariance matrix of the residuals, which depends on i only through its dimension. This general linear mixed model implies that the marginal density function of $\boldsymbol{Y_i}$ is an n_i-dimensional normal distribution with mean $E(\boldsymbol{Y_i}) = \boldsymbol{X_i\beta}$, and with variance covariance matrix $\text{Cov}(\boldsymbol{Y_i}) = \boldsymbol{Z_iDZ_i}' + \boldsymbol{R_i}$. For the estimation of within- and between- subject correlation coefficients, we assume random intercepts only and a Kronecker product structure for $\boldsymbol{R_i}$ ($\boldsymbol{R_i} = \dim_{qmi}(\boldsymbol{V} \otimes \boldsymbol{\Sigma})$), such that the q x q block diagonal of $\boldsymbol{R_i}$ provides estimates of the within-subject correlations between responses and $\boldsymbol{D}$ provides estimates of the between-subject correlation between responses. With this specification of $\boldsymbol{R_i}$, it is assumed that $\boldsymbol{V}$, which is a p x p correlation matrix of the repeated measures on a given response, is the same for all response variables. It is also assumed that $\boldsymbol{\Sigma}$, which is a q x q dimensional positive definite matrix representing the variance covariance matrix between the responses at a given time point, is the same for all time points. SAS PROC MIXED (13) allows for the specification of three different error structures for $\boldsymbol{V}$, namely unstructured, compound symmetry, and first-order autoregressive. Lastly, the overall correlation coefficient can be estimated from the overall covariance matrix $\boldsymbol{\Omega}$.

The bivariate or multivariate LMM can also be specified in such a way that correlations indicating the association of the evolutions can be estimated instead of the above described within- and between-subject correlations. This is accomplished by allowing both random intercepts and random slopes in the model; the correlation of the evolutions is then simply the correlation between the random slopes. Additionally, by not assuming independence of the observations given the random effects, correlations between the residual errors can also be estimated by specifying a Kronecker product structure for $\boldsymbol{R_i}$. With this specification, $\boldsymbol{R_i}$ yields estimates of the residual within-subject correlations that remain after accounting for the subject-specific time trends in each response.

2.2 Univariate linear mixed modeling strategy

As noted, bivariate and multivariate LMMs are preferred over univariate LMMs for the estimation of correlations between responses measured repeatedly over time. However, univariate LMMs are frequently used to estimate such associations, and it is of interest to note how the results of these models differ from those of bivariate and multivariate LMMs. Within- and between-subject correlation coefficients between two responses can be calculated from univariate LMMs; however, the inclusion of more than one response as a covariate yields estimates of association that are typically not the target quantities of interest. To compute within- and between-subject correlation coefficients, the following univariate LMM can be fit (14, 15):

$$Y_{ij} = \alpha_j + \beta_{Wj}(X_{ij} - \bar{X}_j) + \beta_B \bar{X}_{j\,+} \varepsilon_{ij}$$
$$\alpha_j = \alpha + u_{0j}\,,\ \beta_{wj} = \beta_w + u_{1j}$$
$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_B^2 & \sigma_{B,\beta_W} \\ \sigma_{B,\beta_W} & \sigma_{\beta_W}^2 \end{pmatrix}\right],\ \varepsilon_{ij} \sim \mathrm{N}(0,\ \sigma_\varepsilon^2)$$

where $j = 1 \ldots J$ indexes individual subjects; $i = 1 \ldots n_j$ indexes repeated measurements within subject j, and $\bar{X}_j$ is subject j's mean X over all measurements. In this model, β_B reflects the between-subject association between Y and X and β_W reflects the within-subject association (15). In general, when only one term for the association between Y and X is included in the model, its coefficient reflects a weighted average of the within- and between-subject associations, and when the two differ, it is best to specify a model with separate within- and between-subject terms (15). By fitting additional univariate LMMs, it is possible to estimate within- and between-subject coefficients (14) as follows:

$$\rho_{YX,between} = \hat{\beta}_B \sqrt{\frac{\hat{\sigma}_{B,X}^2}{\hat{\sigma}_{B,Y}^2}}$$

$$\rho_{YX,within} = \hat{\beta}_W \sqrt{\frac{\hat{\sigma}_{W,X}^2}{\hat{\sigma}_{W,Y}^2}}$$

where the variance components for X and Y are estimated in the following LMMs:

$$A_{ij} = \alpha_{j,A} + \varepsilon_{Aij}$$
$$\alpha_{j,A} = \alpha_A + u_{0j,A}$$
$$u_{0j,A} \sim N(0, \sigma_{B,A}^2)\ ,\ \varepsilon_{0j,A} \sim N(0, \sigma_{W,A}^2)$$

where $A = Y, X$. The overall correlation coefficient can be estimated as

$$\rho_{YX,total} = \hat{\beta}_{total} \sqrt{\left[\frac{\hat{\sigma}_{B,X}^2 + \hat{\sigma}_{W,X}^2}{\hat{\sigma}_{B,Y}^2 + \hat{\sigma}_{W,Y}^2}\right]}$$

where $\widehat{\boldsymbol{\beta}}_{\boldsymbol{total}}$ can be estimated from a univariate LMM for Y in which only one coefficient is estimated for the association between Y and X (14).

3. Example Data

We illustrate the estimation of the various types of correlations between longitudinal responses using data from a prospective observational study of 23 infants and children treated with extracorporeal membrane oxygenation (ECMO). Animal studies have demonstrated that cell-free plasma oxyhemoglobin (OxyHgb), resulting from the rupture of red blood cells within the blood vessel (intravascular hemolysis), rapidly reacts with intravascular nitric oxide (NO), leading to endothelial dysfunction and end-organ dysfunction (16). Patients on ECMO provide a setting in which to evaluate this phenomenon in humans because intravascular hemolysis occurs secondary to the sheer forces associated with the pump and circuit of these life-support machines. In this study, serial blood samples were collected before, during, and after ECMO in these patients. Plasma was isolated and evaluated for cell-free hemoglobin (Hgb) levels, hemoglobin species (oxyhemoglobin, methemoglobin), and plasma NO consumption using absorbance spectrophotometry, spectral deconvolution and NO chemiluminescence assays, respectively. For this analysis, we used data on Hgb, OxyHgb, and NO consumption from samples that were collected at 2, 4, and 6 hours after the start of ECMO, daily for the remainder of the child's first week on ECMO, and at least one time/week during the remainder of the child's time on ECMO. Only data from the first 21 days of each ECMO course were included in analyses. This cut-off was chosen because only 2 patients had ECMO courses longer than 21 days and we wished to minimize the loss of data and any undue influence these patients might have had on the results while maximizing the likelihood of convergence of the bivariate and multivariate models. For the bivariate and multivariate LMMs, because the Kronecker product covariance structures available in SAS PROC MIXED require a categorical time variable, the time points were grouped into the minimum number of periods such that no patient had more than one measurement per period, yielding 19 time periods.

4. Results

Figure 1 shows the trends in all factors over time, averaged over all patients and modelled using LMMs with time effects fit using restricted cubic splines with a knot at each decile. Figure 2 shows individual patients' trends in each factor over time for 4 randomly selected patients. These figures visually suggest the presence of positive overall and within-subject correlations between NO Consumption and both Hgb and OxyHgb.

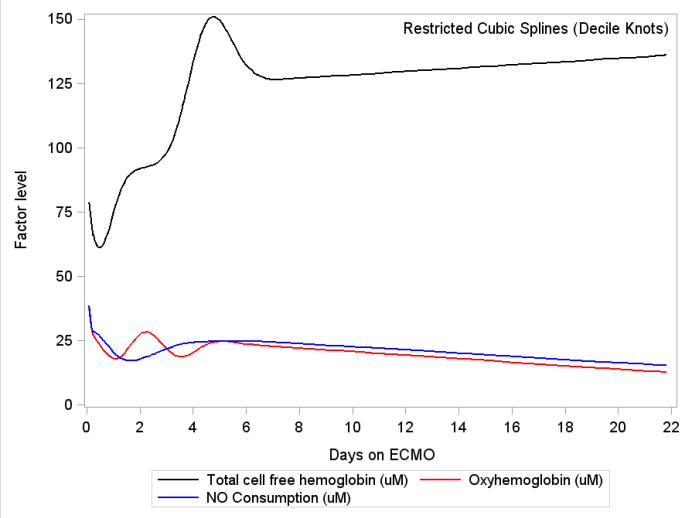

Figure 1: Average trends in all factors over time

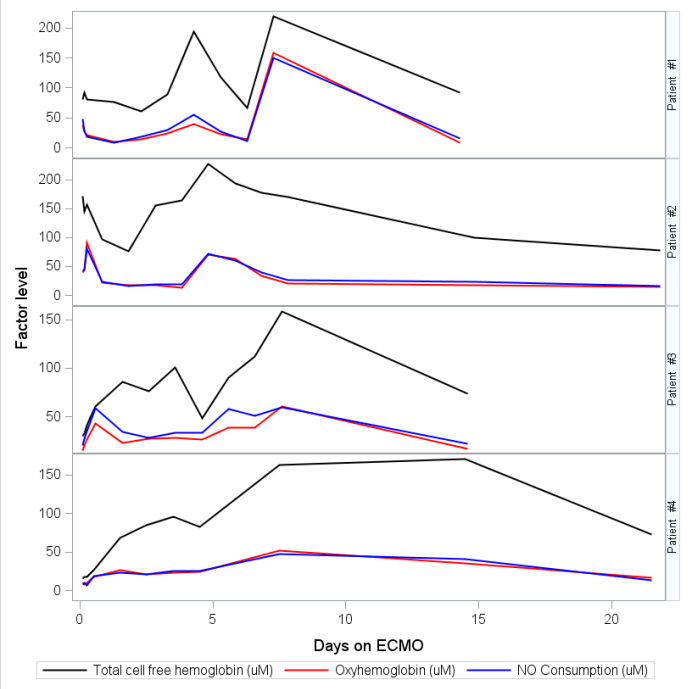

Figure 2: Individual patient trends in all factors over time for 4 randomly selected patients

Due to the skewed distribution of the responses, each was log transformed prior to fitting the models used to calculate correlations. The default restricted error maximum likelihood method (REML) and the Kenward and Roger method for computing the denominator degrees of freedom of the fixed effects were used. The V and Vcorr options were used to output the $\boldsymbol{\Omega}$ variance-covariance matrix and its corresponding correlation matrix. The G and Gcorr options were used to output the $\boldsymbol{D}$ variance-covariance matrix and its corresponding correlation matrix. Finally, the R and Rcorr options were used to output the $\boldsymbol{R_i}$ variance-covariance matrix and its corresponding correlation matrix. The following SAS code was used to fit the bivariate and multivariate LMMs for the estimation of within- and between-subject correlation coefficients. In this example, a Kronecker product variance covariance structure for $\boldsymbol{R_i}$ with $\boldsymbol{V}$ having a compound

symmetry structure was specified. For the bivariate models, "where" statements were used to restrict the analyses to only the two variables of interest.

```
proc mixed data=ecmo_data_long;
 class patient mvar replicate;
 model logresponse = mvar / solution ddfm=kr;
 random mvar / type=un subject=patient V Vcorr G Gcorr;
 repeated mvar replicate / type=un@cs subject=patient R Rcorr;
run;
```

These bivariate and multivariate LMMs yielded the overall, within- and between-subject correlation coefficients shown in Table 1.

Table 1: Overall, within-, and between-subject correlation coefficients from bivariate and multivariate linear mixed models fit using un@cs and un@ar(1) variance covariance structures for $\boldsymbol{R_i}$

Structure of *V* and type of correlation	**Total cell-free hemoglobin and NO Consumption**				**Oxyhemoglobin and NO Consumption**			
	Bivariate model		Multivariate model		Bivariate model		Multivariate model	
	r	P	r	P	r	P	r	P
CS								
Overall	0.37	<.0001	0.37	<.0001	0.62	<.0001	0.62	<.0001
Between-subject	0.22	0.75	0.22	0.25	0.88	0.06	0.88	0.054
Within-subject	0.41	<.0001	0.42	<.0001	0.56	<.0001	0.57	<.0001
AR(1)								
Overall	0.44	<.0001	0.42	<.0001	0.63	<.0001	0.64	<.0001
Between-subject	Inestimable		-0.08	1.0	0.94	0.002	1.0	0.07
Within-subject	0.52	<.0001	0.50	<.0001	0.57	<.0001	0.60	<.0001

These results suggest, as we could predict from the plots, that the within-subject correlations between NO consumption and both total cell-free hemoglobin and oxyhemoglobin were highly statistically significant. Due to the small number of patients and the greater variability within-subjects than between-subjects in the factors, the between-subject correlation coefficients were found to be nonsignificant in all models except the bivariate model for oxyhemoglobin and NO consumption with $\boldsymbol{R_i}$ having an un@ar(1) structure. In the bivariate model for total cell-free hemoglobin and NO consumption with $\boldsymbol{R_i}$ having an un@ar(1) structure, the between-subject correlation coefficient was inestimable because the between-subject variance estimate in $\boldsymbol{D}$ was zero for NO consumption. In addition, the bivariate and multivariate LMMs with $\boldsymbol{R_i}$ specified as un@un did not converge. Likelihood ratio tests were used to test the statistical significance of each correlation. For example, the following code was used to fit a model in which the within-subject correlation coefficient was set equal to zero; the difference between twice the log-likelihood of this model and that of the model with this correlation allowed to vary was compared to a chi-squared distribution with one degree of freedom to yield the p-value shown in Table 1 for this correlation.

```
proc mixed data=ecmo_data_long;
 class patient mvar replicate;
 model logresponse = mvar / solution ddfm=kr;
 random mvar / type=un subject=patient V Vcorr G Gcorr;
 repeated mvar replicate / type=un@cs subject=patient R Rcorr;
 parms (1) (0) (1) (1) (0) (1) (0.5) / hold= 5;
run;
```

The following SAS code was used to fit multivariate LMMs for the estimation of the correlations of the evolutions in the responses over time. In this example, a Kronecker product variance covariance structure was specified for $\boldsymbol{R_i}$ with $\boldsymbol{V}$ having a compound symmetry structure; this yielded estimates of the residual within-subject correlations that remained after accounting for the subject-specific linear changes over time in each factor. Continuous time on ECMO was treated as a separate random effect for each response in this model.

```
proc mixed data=ecmo_data_long;
 class patient mvar replicate;
 model logresponse = mvar mvar*days_on_ECMO / solution ddfm=kr;
 random mvar mvar*days_on_ECMO / type=un subject=patient V Vcorr G Gcorr;
 repeated mvar replicate / type=un@cs subject=patient R Rcorr;
run;
```

These multivariate LMMs yielded the correlations of the evolutions and the residual within-subject correlations shown in Table 2.

Table 2: Correlations between linear trends over time and residual within-subject correlations from multivariate linear mixed models

	$\boldsymbol{R_i}$ with un@cs structure		$\boldsymbol{R_i}$ with un@ar(1) structure	
	r	P	r	P
Correlation of the evolutions				
Hemoglobin and NO Consumption	0.25	0.58	0.07	1.00
Oxyhemoglobin and NO Consumption	0.28	0.58	0.47	0.24
Residual within-subject correlations				
Hemoglobin and NO Consumption	0.53	<.0001	0.51	<.0001
Oxyhemoglobin and NO Consumption	0.63	<.0001	0.63	<.0001

These models show that the residual within-subject correlations are much stronger than the correlations of the evolutions, which is not surprising given the clearly nonlinear changes in all factors over time. If a better fitting functional form had been chosen for the evolutions in the responses over time, their correlations likely would have been stronger. However, due to the heterogeneity in all factors' changes over time within subjects, we preferred to compare the structures available in SAS PROC MIXED for the variance covariance matrix of the residuals rather than to search for an ideal functional form for the fixed time effects.

Finally, univariate LMMs were used to estimate within- and between-subject correlations between total cell-free hemoglobin and NO Consumption using the following SAS code.

```
PROC mixed data= all_ecmo_data;
  class patient;
  model logNO_Cons = logHgbdiff meanlogHgb / s ddfm=kr;
  random intercept logHgbdiff / type=un subject=patient;
run;

PROC mixed data= all_ecmo_data;
  class patient;
  model logNO_Cons =   / s ddfm =kr;
  random intercept / type=un subject=patient;
run;

PROC mixed data= all_ecmo_data;
  class patient;
  model logHgb=   / s ddfm =kr;
  random intercept / type=un subject=patient;
run;
```

The significance of the within- and between-subject effects in the first model indicated the statistical significance of the corresponding correlation coefficients. The second and third models were required to estimate the variance parameters needed to calculate the within- and between-subject correlation coefficients. These variance parameters, along with the parameter estimate for total cell-free hemoglobin as a single effect in a separate univariate LMM, were used to estimate the overall correlation coefficient. Identical models were fitted for NO Consumption with oxyhemoglobin as the predictor. These correlation coefficients are shown in Table 3.

Table 3: Overall, within-, and between-subject correlation coefficients from univariate linear mixed models

Type of correlation	**Total cell-free hemoglobin and NO Consumption**		**Oxyhemoglobin and NO Consumption**	
	r	P	r	P
Overall	0.37	<.0001	0.60	<.0001
Between-subject	0.25	0.28	0.79	<.0001
Within-subject	0.41	<.0001	0.56	<.0001

All point estimates are similar among the univariate, bivariate, and multivariate models. However, the main difference between the correlation coefficients produced by the univariate LMMs and those produced by the bivariate and multivariate LMMs is the greater statistical significance of the between-subject correlation between oxyhemoglobin and NO consumption in the univariate model. This difference is mainly because the univariate model does not account for the variability of the estimate of the subject's average oxyhemoglobin over their entire ECMO course; this is treated as an independent variable in the model and is thus assumed to be measured without error.

4. Conclusions

In this study, we have compared several different types of correlation coefficients that can be estimated from univariate, bivariate, and multivariate linear mixed models for longitudinal data on multiple responses. With such data, the Pearson correlation coefficient is not an appropriate estimator because it ignores the correlation structure of the repeated measures. We showed that the various types of correlations that can be calculated from longitudinal data can be quite different, but that generally the three types of models yield similar point estimates for overall, within-, and between-subject correlation coefficients. It is important to understand the types of correlations that one wishes to estimate and the correct interpretation of those correlations prior to model fitting.

There are several limitations inherent in the models fit in this study. Firstly, the Kronecker product variance covariance structure for $\boldsymbol{R}_i$ requires the restrictive assumption that the intra-response correlations are equal for all responses. Although a fully unstructured variance covariance matrix could be specified for $\boldsymbol{R}_i$, this model would not converge in our dataset and likely would not converge in most datasets with highly unbalanced data. Additionally, all of the Kronecker product variance covariance structures currently available in SAS PROC MIXED best accommodate equally spaced time points. In the future, we would like to investigate structures more appropriate for unequally spaced time points. In this study we did not compare goodness of fit criteria such as Akaike's information criterion (AIC), AIC corrected (AICC), or Bayesian information criterion (BIC) between models. We chose not to examine these criteria because 1) they could not have distinguished the best fitting model of all the models that were fit and 2) the goal of this study was to compare and contrast the different types of correlation coefficients estimable from univariate, bivariate, and multivariate LMMs rather than to find the best fitting model. Finally, we did not discuss all of the diagnostic analyses that should typically be incorporated into any analysis of repeated measures data. However, we did assess the normality assumption of the outcomes and residuals as well as the effects of influential points and outliers. In conclusion, multivariate longitudinal data provide a unique opportunity to examine the joint progression of multiple responses over time. Since several distinct types of correlations can be estimated from such data, it is important to understand which correlation is of interest prior to model fitting. Subsequently, the best fitting bivariate or multivariate linear mixed model that answers this question should be identified.

References

1. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 1--Correlation within subjects. BMJ. 1995;310(6977):446.
2. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 2--Correlation between subjects. BMJ. 1995;310(6980):633.
3. Lam M, Webb KA, O'Donnell DE. Correlation between two variables in repeated measures. Joint Statistical Meeting: American Statistical Association; 1999.
4. Hamlett A, Ryan L, Wolfinger R. On the use of PROC MIXED to Estimate Correlation in teh Presence of Repeated Measures. 29th SAS Users Group International Conference; Montreal, Canada. Cary, NC: SAS Institute Inc.; 2004.

5. Roy A. Estimating correlation coefficient between two variables with repeated observations using mixed effects model. Biometrical Journal. 2006;48(2):286-301.
6. Thiebaut R, Jacqmin-Gadda H, Chene G, Leport C, Commenges D. Bivariate linear mixed models using SAS proc MIXED. Comput Methods Programs Biomed. 2002;69(3):249-56.
7. Fieuws S, Verbeke G. Joint modelling of multivariate longitudinal profiles: pitfalls of the random-effects approach. Stat Med. 2004;23(20):3093-104.
8. Fieuws S, Verbeke G. Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. Biometrics. 2006;62(2):424-31.
9. Hosmer DW, Lemeshow S, May S. Applied survival analysis : regression modeling of time-to-event data. 2nd ed. Hoboken, N.J.: Wiley-Interscience; 2008. xiii, 392 p. p.
10. Bandyopadhyay S, Ganguli B, Chatterjee A. A review of multivariate longitudinal data analysis. Stat Methods Med Res. 2011;20(4):299-330.
11. Rochon J. Analyzing bivariate repeated measures for discrete and continuous outcome variables. Biometrics. 1996;52(2):740-50.
12. Gray SM, Brookmeyer R. Estimating a treatment effect from multidimensional longitudinal data. Biometrics. 1998;54(3):976-88.
13. SAS Institute Inc. 2011. SAS/STAT® 9.3 User's Guide. Cary, NC: SAS Institute Inc.
14. Ferrari P, Al-Delaimy WK, Slimani N, Boshuizen HC, Roddam A, Orfanos P, et al. An approach to estimate between- and within-group correlation coefficients in multicenter studies: plasma carotenoids as biomarkers of intake of fruits and vegetables. Am J Epidemiol. 2005;162(6):591-8.
15. Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis. 2nd ed. Hoboken, N.J.: Wiley; 2011. xxv, 701 p. p.
16. Minneci PC, Deans KJ, Zhi H, Yuen PS, Star RA, Banks SM, et al. Hemolysis-associated endothelial dysfunction mediated by accelerated NO inactivation by decompartmentalized oxyhemoglobin. J Clin Invest. 2005;115(12):3409-17.