Estimating Subject-Specific Rates of Change from Longitudinal Data

Christopher H. Morrell[1,2], Larry J. Brant[1], Majd AlGhatrif[2], Edward G. Lakatta[2]
[1]Mathematics and Statistics Department, Loyola University Maryland, Baltimore, MD
[2]Laboratory of Cardiovascular Science, NIA IRP, NIH, Baltimore, MD

Abstract
In a longitudinal study it is often of interest to study individual differences by estimating the rate of change in the response variable for each subject. If the data is modeled using a linear mixed-effects model that includes fixed and random effects for time, a subject specific rate of change can be obtained as the derivative of the model function with respect to time and can be estimated using the fixed and random effects for time. However, due to shrinkage these subject-specific rates can appear to have much less variability than might be anticipated to be clinically useful. Conversely, if regression models are fit to each subject's data the resulting rates have too much variability due to the small number of observations for each subject as well as the within subject error. The goal of this talk is to combine these two estimates to obtain a better estimate of subject specific rates of change. This is accomplished by taking a weighted average of the two estimates using weights proportional to the reciprocal of the variances. Various approaches are investigated on how to implement this strategy and are illustrated using longitudinal data from the Baltimore Longitudinal Study on Aging.

Key Words: Mixed-Effects Models; Shrinkage; Inverse Variance Weighting

1. Introduction

Data from longitudinal studies are used to describe trends over time in the subjects being studied. In addition to obtaining average trajectories and rates of change for all subjects in the study, it is also desirable to understand individual variability or differences by obtaining subject-specific trajectories and rates of change.

Linear mixed-effects (LME) models have become a standard statistical model for analyzing the repeated-measures data that derives from longitudinal studies (Laird and Ware, 1982; Morrell, Pearson, and Brant, 1997; Verbeke and Molenberghs, 2000; Gueorguieva and Krystal, 2004; Morrell, Brant, and Ferrucci, 2009). These models handle unbalanced data while including any relevant explanatory variables that need to be taken into account. The models contain both fixed- and random-effects. The fixed effects allow for the estimation of population average trajectories and rates of change while the inclusion of the subject-specific random effects enables the variability in these quantities among subjects to be investigated. The estimates of the subject-specific random effects are shrinkage estimators as the estimates for each individual "borrow strength" from the other subjects in the study to obtain "better" subject-specific estimates resulting in subject-specific estimates that are shrunken towards to overall mean. The amount of shrinkage depends on the amount of data for the subject and the relative between and within subject error variances. Singer and Willett (2003) discuss the estimation of subject specific rates of change. They consider three possibilities: population average rates of change, estimates obtained from subject-specific regressions, and the shrinkage estimates described above.

Due to the shrinkage, the subject specific rates of change can appear to have too little variability compared to what one might expect to be of use in a clinical setting. On the other hand, if linear regression models are fit to each subject's data individually, there will be too much variability due to the small number of observations for each subject. The goal of this paper is to propose and study an alternative strategy to obtaining subject-specific rates of change. The proposed estimate is a weighted average of the LME estimate and the regression estimate. The weights used are proportional to the inverse of the variances of each of the estimates (Hartung, Knapp and Sinha, 2008). A number of strategies are compared and evaluated on how to combine the two estimates.

The remainder of the paper is organized as follows. Section 2 describes the linear mixed effects model and its rate of change. Section three indicates how to combine the estimates of rates of change from the LME model with estimates from subject-specific regressions. Section 4 uses an example to illustrate the proposed approaches and section 5 draws conclusions from the study.

2. Linear Mixed-Effects Model

For subject *i*, the model is

$$y_i = X_i\beta + Z_i b_i + \varepsilon_i, i = 1, \ldots, m$$

where y_i is the $n_i \times 1$ vector of the response variable for subject *i*, X_i is the design matrix of fixed effects covariates, β is the $p \times 1$ vector of regression parameters for the fixed-effects, Z_i is the design matrix of random effects, b_i is the $q \times 1$ vector of subject-specific random effects, and ε_i is the error term. Further, $b_i \sim N(0, D), \varepsilon_i \sim N(0, \sigma^2 I)$, and b_i and ε_i are independent. Once the fixed-effects parameter vector, β, is estimated, the individual random effects are calculated as:

$$\hat{b}_i = DZ_i^T V_i^{-1}(y_i - X_i\hat{\beta})$$

where $V_i = Z_i D Z_i^T + \sigma^2 I$. As is well known, these estimates are shrinkage estimates and they "borrow strength" from the rest of the data to obtain "better" estimates than obtaining estimates from each individual's data separately.

However, these individual estimates of rates of change may be shrunken too much (in that they have too little variability relative to what one might expect, even accounting for the variation in level and rate of change among individuals). To attempt to ameliorate this problem, we propose to estimate the individual rates of change by a weighted average of two estimates: the estimates from the LME model and the estimates from a subject-specific regression (this will, of course, require the subject to have at least two observations).

To set the stage, let the LME model be:

$$\hat{y}_{i,LME} = \hat{\beta}_0 + \hat{b}_{i0} + (\hat{\beta}_1 + \hat{b}_{i1})Time.$$

This model has only a single Time term. For the purposes of this discussion, without loss of generality, we can ignore other terms unrelated to Time in the model as they will not affect the rate of change with respect to Time. The estimated rate of change for subject *i* is:

$$\widehat{Rate}_{i,LME} = (\hat{\beta}_1 + \hat{b}_{i1}). \quad (1a)$$

However, the LME model may contain an interaction term of the entry age (EAge) with Time. This allows the rate of change to increase/decrease linearly with the EAge. For example, suppose

$$\hat{y}_{i,LME} = \hat{\beta}_0 + \hat{b}_{i0} + (\hat{\beta}_1 + \hat{b}_{i1})Time + \hat{\beta}_2 EAge + \hat{\beta}_3 EAge \times Time.$$

In this case the estimated rate of change for subject *i* is:

$$\widehat{Rate}_{i,LME} = \left(\hat{\beta}_1 + \hat{b}_{i1}\right) + \hat{\beta}_3 EAge. \qquad (1b)$$

3.1 Combining Rate of Change Estimates

Let the usual ordinary least squares regression model for the data from only subject *i* be:

$$\hat{y}_{i,OLS} = \hat{\theta}_0 + \hat{\theta}_{1i} Time$$

with rate of change for subject *i* given by

$$\widehat{Rate}_{i,OLS} = \hat{\theta}_{1i} \qquad (2)$$

We propose combining the two estimated rates of change ((1a) or (1b) with (2)) by using the inverse of the variances of each of the estimates (Hartung, Knapp and Sinha, 2008). The LME and regression rate estimates have standard errors that can be obtained from the LME (proc mixed) and Regression outputs in SAS. The standard error of the rate estimate from the LME model (1a) is given by

$$SE\left(\widehat{Rate}_{i,\,LME}\right) = \sqrt{SE\left(\hat{\beta}_1\right)^2 + SE\left(\hat{b}_{i1}\right)^2 + 2 \times Cov(\hat{\beta}_1, \hat{b}_{i1})}.$$

The standard error of (1b) would include terms involving the covariance of the two fixed-effects parameter estimates, $\hat{\beta}_1$ and $\hat{\beta}_3$ with each other and with the estimated random effect, $\hat{b}_{i1}$. To evaluate this expression the covariance between the fixed and random effects of Time is required. It turns out that this covariance is small and can be omitted without much effect on the standard errors. Here we include the covariance term by employing the estimate statement in proc mixed in SAS to obtain the correct standard error for (1a) (see solution posted at http://support.sas.com/kb/37/109.html).

To combine the two estimates, $\widehat{Rate}_{i,LME}$ and $\widehat{Rate}_{i,OLS}$ a weighted average is used:

$$\widehat{Rate}_{i,Comb} = \lambda_i\, \widehat{Rate}_{i,LME} + (1 - \lambda_i)\, \widehat{Rate}_{i,OLS} \qquad (3)$$

where we need to determine the weights, λ_i. A good choice is to use the reciprocal of the variances of the two estimates to obtain the weights as this will minimize the variance of the resulting weighted average (Hartung, Knapp, and Sinha, 2008). Then the estimate with the smaller variance (SE) will be weighted more heavily. That is:

$$\lambda_i = \frac{1/Var(\widehat{Rate}_{i,LME})}{1/Var(\widehat{Rate}_{i,LME}) + 1/Var(\widehat{Rate}_{i,OLS})} \qquad (4)$$

where the variances are the squares of the standard errors (SEs) of the estimates discussed above.

3.2 Simple Example Illustrating the Shrinkage

For purposes of illustration, we use a simple example in which we assume that each subject has two observations at times 0 and t. The model we consider for this example is:

$$y_i = (\beta_1 + b_{i1})T + \varepsilon_i$$

This is a simple regression model through the origin. We assume that $\varepsilon_i \sim N(0, \sigma^2)$ and $b_{i1} \sim N(0,\ \sigma_T^2)$.

Then $X_i = Z_i = \begin{pmatrix} 0 \\ t \end{pmatrix}$. Using the formulae from Laird and Ware (1982) (see below) yields the following.

$$V_i = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} + \begin{pmatrix} 0 \\ t \end{pmatrix} \sigma_T^2 (0 \quad t) = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 + t^2\sigma_T^2 \end{pmatrix}$$

and

$$W_i = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/(\sigma^2 + t^2\sigma_T^2) \end{pmatrix}.$$

So

$$\hat{\beta} = \left(\sum_{i=1}^{m} \frac{t^2}{\sigma^2 + t^2\sigma_T^2} \right)^{-1} \sum_{i=1}^{m} \frac{t y_{i2}}{\sigma^2 + t^2\sigma_T^2} = \left(\frac{\sigma^2 + t^2\sigma_T^2}{mt^2} \right) \left(\frac{t}{\sigma^2 + t^2\sigma_T^2} \right) \sum_{i=1}^{m} y_{i2} = \frac{1}{t} \bar{y}_2.$$

The estimated average slope is just the slope from the origin to the mean at the second time point. Since the line is being forced through the origin, the observation at the first time point plays no part in the estimation of the slope. Next

$$\begin{aligned} \hat{b}_i = \ & \sigma_T^2 (0 \quad t) \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & 1/(\sigma^2 + t^2\sigma_T^2) \end{pmatrix} \left(\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} - \begin{pmatrix} 0 \\ t \end{pmatrix} \frac{1}{t} \bar{y}_2 \right) \\ & = \sigma_T^2 \begin{pmatrix} 0 & \frac{t}{\sigma^2 + t^2\sigma_T^2} \end{pmatrix} \begin{pmatrix} y_{i1} \\ y_{i2} - \bar{y}_2 \end{pmatrix} = \frac{t\sigma_T^2}{\sigma^2 + t^2\sigma_T^2} (y_{i2} - \bar{y}_2) \\ & = \frac{T}{\frac{\sigma^2}{\sigma_T^2} + t^2} (y_{i2} - \bar{y}_2) \,. \end{aligned}$$

Consequently, the estimated slope for the ith subject from the mixed-effects model is:

$$\hat{\beta} + \hat{b}_i = \frac{\bar{y}_2}{t} + \frac{t}{\frac{\sigma^2}{\sigma_T^2} + t^2} (y_{i2} - \bar{y}_2) = \left(\frac{1}{\frac{\sigma^2}{t^2\sigma_T^2} + 1} \right) \frac{y_{i2}}{t} + \left(1 - \frac{1}{\frac{\sigma^2}{t^2\sigma_T^2} + 1} \right) \frac{\bar{y}_2}{t}.$$

This is a weighted average of the slope from the origin for the ith subject, $\frac{y_{i2}}{t}$, and the mean slope, $\frac{\bar{y}_2}{t} = \hat{\beta}$. When the error variance, σ^2, is large relative to the between subject variance in the slopes, σ_T^2, then the observed data gets less weight and the overall mean slope gets more weight – and conversely.

Now, when we take a weighted average of this estimate and the LS estimate from the data, we obtain

$$\begin{aligned} & \lambda \left(\left(\frac{1}{\frac{\sigma^2}{t^2\sigma_T^2} + 1} \right) \frac{y_{i2}}{T} + \left(1 - \frac{1}{\frac{\sigma^2}{t^2\sigma_T^2} + 1} \right) \frac{\bar{y}_2}{t} \right) + (1 - \lambda) \frac{y_{i2}}{t} \\ & \qquad = \left(\lambda \left(\frac{1}{\frac{\sigma^2}{t^2\sigma_T^2} + 1} \right) + (1 - \lambda) \right) \frac{y_{i2}}{t} + \lambda \left(1 - \frac{1}{\frac{\sigma^2}{t^2\sigma_T^2} + 1} \right) \frac{\bar{y}_2}{t} \end{aligned}$$

so that the observed data is weighted even more heavily and there will be less shrinkage towards the overall mean slope.

3.3 Combining Rate of Change Estimates - Continued

There are some limitations to the approach presented above in §3.1. For the regression approach a subject must have at least two observations to obtain an estimate of their rate of change and at least three observations to be able to obtain an estimate of the standard error. In addition, if a subject has three or more observations that are exactly linear, the standard error will be 0. Consequently, the approach, as stated, cannot obtain combined estimates for subjects with only two observations or with more observations that are colinear. We seek to address this limitation in a number of ways. There are three main categories of approaches to obtaining the weights, each with two sub-approaches. The two sub-approaches address (a) working directly with the λ_i (4) and (b) working with each of the standard errors on the right-hand side of (4). The main approaches are:

1. (a) Average the weights for all subjects or
   (b) average the two standard errors and use these averages to compute the weights as in (4).
   In this case all subjects will have the same weight, regardless of the number of observations available for each subject.
2. The weights are associated with the amount of data for each subject. Compute the mean weight and mean of the standard errors for each number of observations, n_i. Then fit linear regression models:
   (a) to the mean weight vs. number of observations or
   (b) separate regressions of the means of the standard errors vs. number of observations.
   Use predictions from the regressions to obtain the weights. Here the weight usually declines with number of observations so that the weight given to the regression estimate (2) increases as the amount of data for the subject increases.
3. Model:
   (a) the individual weights as a function of number of observations or
   (b) each of the individual standard errors as a function of number of observations.
   Use predictions from the regressions to obtain the weights. As in 2, above, the regression estimate usually gets more weight as the amount of data for the subject increases.

These three approaches allow for the computation of a combined rate in cases that were not possible using (4), i.e when a subject has only two observations or when the subject's data is exactly linear. Including the LME rate, equations (1a) or (1b), the subject-specific regression rates (2), and the weighted version given in (4), this results in 9 different estimates to consider and compare.

4. Example

The example uses data from the Baltimore Longitudinal Study of Aging (BLSA) (Shock et. al, 1984). The BLSA is an ongoing multidisciplinary longitudinal study that began in 1958. Participants return to Baltimore at about 2-year intervals and undergo testing and measurement on numerous factors. Here we apply the methods presented above to glucose.

The data set consists of 3202 observations on 1266 participants (an average of 2.5 observations per participant, maximum of 9 observations). The fixed-effects parameter estimates in the LME model for glucose are given in Table 1. The final model includes random effects for intercept, Time, and Time^2. Here we are considering the rate of change in glucose at the initial visit. Consequently, the Time^2 term is eliminated from the estimate of the rate of change. The average rate of change at the initial visit is given by -0.4504. However, the random effect variance is $Var(b_{i1}) = 4.2484$. This suggests that there will be substantial between subject variability in the subject specific rates of change and while, on average, there is a 0.45 initial decline in glucose per year, we will expect many subjects to exhibit increases in their glucose level. Figure 1 compares the estimated rates of change computed from the LME model using (1a) with the rates computed by linear regression. The regression estimates exhibit many more extreme outliers and a variability about twice as large ($IQR_{Regression} = 1.81$ vs. $IQR_{LME} = 0.95$). Next, equation (4) is used to obtain the weights which are applied to obtain the combined estimate.

Table 1. Fixed-effects terms in the final LME model for glucose.

Effect	Estimate	Standard Error	t Value	Pr > |t|
Intercept	60.6485	6.1366	9.88	<.0001
EAge	0.9849	0.2022	4.87	<.0001
EAge^2	-0.00767	0.001623	-4.73	<.0001
Time	-0.4504	0.1313	-3.43	0.0006
Time^2	0.01702	0.009079	1.88	0.0613
Gender	5.2408	0.8265	6.34	<.0001

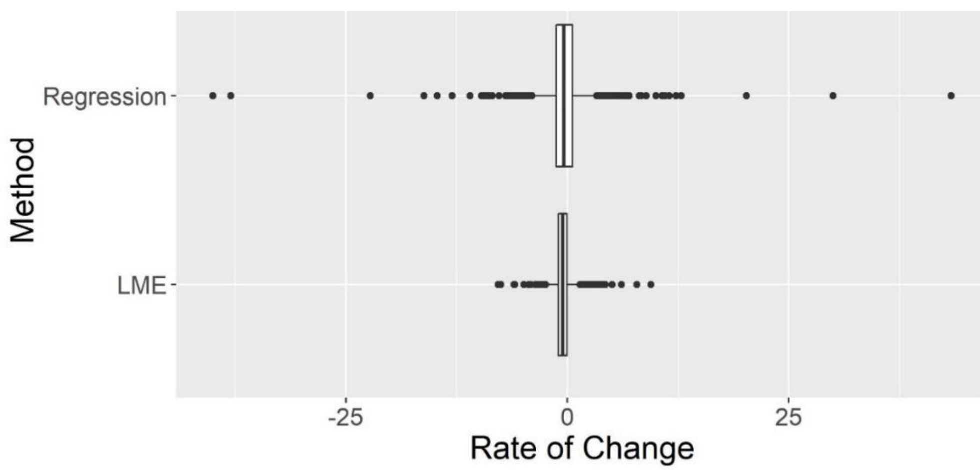

Figure 1. Comparative boxplots of subject-specific rates of change of glucose at initial visit.

Approach 1: Average of weights or SEs.

In this approach, all individuals have the same weight regardless of their number of visits. In addition, subjects with 2 visits or more visits for whom we could not previously obtain a weight, now will be included. The mean of all weights computed by (4) is 0.267. When the averages of the standard errors are used to obtain a common “average” weight for all subjects, the result is 0.307 (mean(SE_{LME}) = 1.686 and mean($SE_{Regression}$) = 1.122). Counterintuitively, note that while the regression estimates of rates of change are more variable over the subjects in the study than the LME estimates, the standard errors of these estimates of the rates of change tend to be smaller, on average, for the regression estimates of the rates of change than for the LME estimates of individual rates of change. Consequently, more weight is given to the regression estimates than to the LME estimates. Averaging the individual SEs leads to more weight being given to the LME rates than averaging the individual weights in (4).

Approach 2: Model average of weights or SEs as a function of number of visits.

In this approach all individuals with the same number of visits will have the same weights.

The model for the average of the weights is $\hat{\lambda} = 0.3525 - 0.0222 \times NVisits$ ($R^2 = 0.91$). Since the association of the weights with the number of visits is negative, as expected the weight given to the subjects' regression estimates (2) increases with the number of visits. Note also that in this example all weights will be less than 0.5 resulting in more weight being given to the regression rate estimates than to the LME rates. See Figure 2.

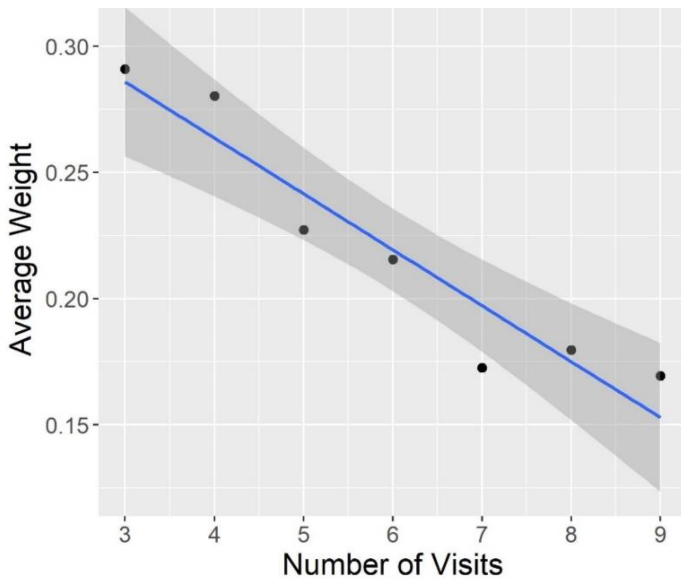

Figure 2. Regression model for the average weight by number of visits.

When the standard errors are plotted against the number of visits curvature is observed. Consequently, quadratic models are fit to the standard errors as a function of number of visits. This gives:

$\widehat{SE}_{LME} = 2.0777 - 0.1558 \times NVisits + 0.00907 \times NVisits^2$ ($R^2 = 0.995$) and

$\widehat{SE}_{Regression} = 2.111 - 0.3255 \times NVisits + 0.01762 \times NVisits^2$ ($R^2 = 0.977$)

Figure 3 provides a comparison. As above, modeling the average SEs rather than the actual weights, themselves, leads to more weight being given to the LME rates (as opposed to modeling the weights, themselves) except when the sample size exceeds 7. From about 6 visits, the two approaches give almost identical results.

Approach 3: Model individual weights or SEs as a function of number of visits.

The model using each individual's weights is $\hat{\lambda} = 0.3745 - 0.0267 \times NVisits$ ($R^2 = 0.017$, P = 0.0028). This is very similar to modeling the average weights. Modeling the individual SEs gives:

$\widehat{SE}_{LME} = 1.959 - 0.0844 \times NVisits$ ($R^2 = 0.503$, P < 0.0001) and

$\widehat{SE}_{Regression} = 1.713 - 0.149 \times NVisits$ ($R^2 = 0.023$, P = 0.0005).

Figure 3 provides a comparison of the weights from all of the approaches. Modeling each individual's SEs leads to more weight being given to the LME rates except when the sample size exceeds 8. Finally, note that the regression weights using individual weights or SEs are similar to the weights obtained from the averages (Figure 3).

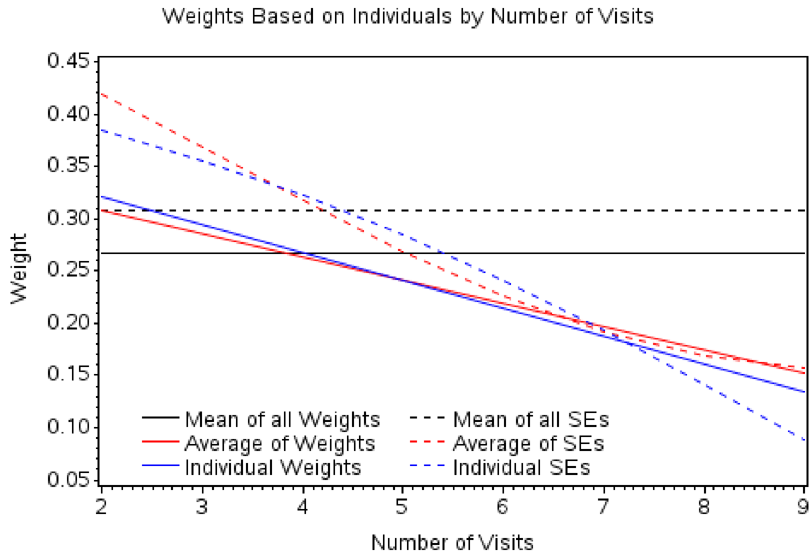

Figure 3. Comparing the weights across the six estimation approaches.

The individual rates from approaches 1 – 3 give very similar results. In particular, their standard deviations all lie between the rates computed from the LME model alone and the original weighted version and the regression estimates (Figure 4, Table 2). Among these, in what follows we restrict our attention to the one with the smallest (2b – regression of average standard errors as a function of sample size) and largest (1a – mean of all weights) standard deviations.

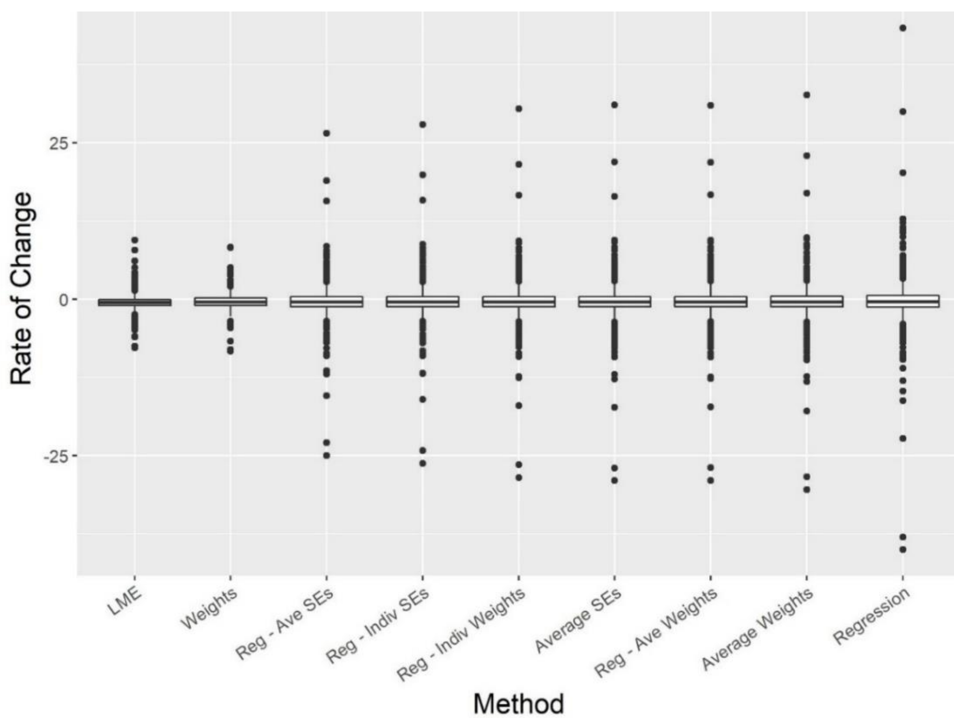

Figure 4. Boxplot Comparisons of the individual rates of change from the nine approaches to estimating the weights.

Table 2. Descriptive statistics of estimated individual rates for the nine approaches to estimating the weights (ordered by standard deviations).

Method		**N**	**Mean**	**Std Dev**
(1a)	**LME**	865	-0.4709	1.1788
(4)	**Weights**	527	-0.3432	1.4632
2b.	**Reg – Ave SEs**	865	-0.3539	2.5775
3b.	**Reg – Indiv SEs**	865	-0.3513	2.6699
3a.	**Reg – Indiv Wgts**	865	-0.3411	2.8553
1b.	**Average SEs**	865	-0.3443	2.8849
2a.	**Reg Ave Wgts**	865	-0.3397	2.8918
1a.	**Average Weights**	865	-0.3370	3.0047
(2)	**Regression**	865	-0.2883	3.8196

Next these rates were summarized separately by the number of repeated observations from which the regression estimate is obtained i.e. the number of observations for each participant. Figure 5 plots the standard deviations as a function of number of visits. The standard deviations converge by about 5 visits. The LME and weighted average approach have consistently lower standard deviations for fewer number of visits.

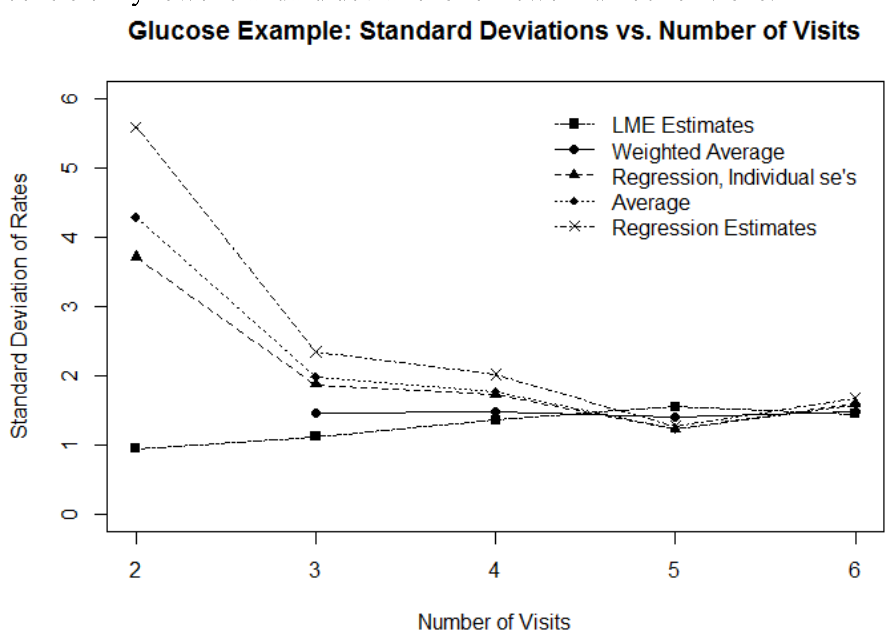

Figure 5. Standard deviation of estimated rates of change by number of visits.

5. Conclusions

The goal of this study was to obtain estimates of the rates of change for each individual in a longitudinal study. Estimates from separate regression will rely on only a few observations for each participant and requires that subjects have at least two observations. In contrast, estimates based on LME models can obtain estimates for all subjects. These estimates borrow strength from other subjects' data to obtain "better" estimates that are shrunken towards the overall average rate of change. However, in applications it appears that these shrinkage estimates may be shrunken too much to be of clinical use. We have proposed a weighted average of the regression and LME estimates of the rates of change. In addition, we have also proposed several versions of the weighted estimate. When the weights are obtained using the reciprocal of the variances of the two estimates (4), the new rates tend to have a standard deviation quite close to that of the rates obtained from the LME model. Next, when the weights are obtained from regression models of the standard errors, the standard deviations of the rates tend to be closer to the rates based on (4) and the LME model. The standard deviation of the rates based on regressions of the actual weights, has still more variability. Finally, using the average of the weights for all subjects tends to have a standard deviation that is closer to the standard deviation of the rates based on the regression models for each individual. In summary the standard deviations of the rates can approximately be ordered as:

LME; (Equation 4); (2b, 3b); (2a, 3a); (1a, b); Individual Regressions.

A researcher will need to choose how much shrinkage is desired. Avoiding too much variability in the final rates suggests that Equation 3 with weights calculated using (4) or regression models based the standard errors might be good choices.

Funding:
This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

References:

Gueorguieva R and Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers in the Archives of General Psychiatry. *Arch Gen Psychiatry* 2004;61:310-317.

Joachim Hartung; Guido Knapp; Bimal K. Sinha (2008). *Statistical meta-analysis with applications*. John Wiley & Sons.

Laird NM, Ware JH. (1982) "Random-effects models for longitudinal data." *Biometrics*. 38(4):963-74.

Morrell CH, Brant LJ, Ferrucci L. (2009). Model Choice Can Obscure Results in Longitudinal Studies. *J Gerontol A Biol Sci Med Sci*, 64A:215–222.

Morrell CH, Pearson JD, and Brant LJ. (1997) "Linear transformations of linear mixed-effects models." *Am Stat,* 51:338-343.

Shock NW, Greulich RC, Andres R, Arenberg D, Costa PT, Lakatta EG, Tobin JD. *Normal human aging: The Baltimore Longitudinal Study of Aging* (NIH Publication No. 84-2450), Washington, DC: U.S. Government Printing Office; 1984.

Singer JD and Willett JB (2003) *Applied Longitudinal Analysis: Modeling Change and Event Occurrence.* Oxford University Press.

Verbeke, G. and Molenberghs, G., (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer.