Model Choice Can Obscure Results in Longitudinal Studies

Christopher H. Morrell,^{1,2} Larry J. Brant,¹ and Luigi Ferrucci³

¹Gerontology Research Center, National Institute on Aging, Baltimore, Maryland.

²Mathematical Sciences Department, Loyola College in Maryland, Baltimore.

³Clinical Research Branch, National Institute on Aging, Harbor Hospital, Baltimore, Maryland.

Background. This article examines how different parameterizations of age and time in modeling observational longitudinal data can affect results.

Methods. When individuals of different ages at study entry are considered, it becomes necessary to distinguish between longitudinal and cross-sectional differences to overcome possible selection biases.

Results. Various models were fitted using data from longitudinal studies with participants with different ages and different follow-up lengths. Decomposing age into two components—age at entry into the study (first age) and the longitudinal follow-up (time) compared with considering age alone—leads to different conclusions.

Conclusions. In general, models using both first age and time terms performed better, and these terms are usually necessary to correctly analyze longitudinal data.

Key Words: Mixed-effects models—Multilevel modeling—Observational study—Recruitment bias—Regression.

EPIDEMIOLOGICAL studies that involve repeated observations from the same individuals over a period of years have provided important information on the etiology, pathophysiology, and consequences of chronic diseases (1). Longitudinal studies are the best method to study individual change (longitudinal differences) and to identify factors associated with that change (2) while still allowing comparison between individuals of different ages (cross-sectional differences). Modeling the average response as a function of age without considering differences in age at study entry assumes that cross-sectional and longitudinal effects of aging are equivalent. On the contrary, it is known that the cross-sectional effects of age at study entry and changes over time are quite different thus requiring special attention in the modeling and interpretation of results.

Schaie (3) discusses how the longitudinal design better separates age, time, and cohort effects. Analyses of repeated measures are often accomplished using hierarchical or multi-level models where age is a Level 1 variable and time is a Level 2 variable. However, in unbalanced longitudinal studies with participants over wide age range, a model should be used that can simultaneously describe cross-sectional and longitudinal patterns. These models will typically require careful parameterization of interactions and higher order terms to adequately fit the data and perhaps handle cohort effects.

Mixed-effects and random-effects models have contributed greatly to the study of longitudinal patterns of change (4–7). In mixed-effects models both individual patterns of change (individual random effects) and population averages (fixed effects) can be estimated. The correlation between the repeated measures is accounted for through the variance-

covariance structure of the random effects. Unlike traditional repeated measures analyses (8) mixed-effects models use all available data, provided data are missing at random (9).

Different approaches have been used to model longitudinally the relationship of age and follow-up time with a response variable. Alfaro-Acha and colleagues (10) in a 7-year prospective study of handgrip strength and cognitive decline in Mexican Americans aged 65 and older used a mixed-effects model including baseline age and time. However, no interaction terms involving baseline age and time were reported, assuming that the longitudinal changes in cognitive decline are not influenced by the individual's starting age. To study hip fracture in female patients aged 65 and older participating in the Baltimore Hip Studies followed up for 12 months, Miller and colleagues (11,12) used baseline age in their generalized estimating equations (GEE) analyses and treated time as a categorical variable. Blaum and colleagues (13) examined the metabolic syndrome in Mexican Americans aged 60–98 years over a 3-year follow-up and modeled their longitudinal data using GEE with both age as a covariate as well as a Time $\times$ Age interaction. Park and Lee (14) considered only age in modeling the occurrence of urinary incontinence during a 2-year period in community-dwelling elderly. In an analysis of data from HIV patients, Edwards and colleagues (15) evaluated longitudinal changes in viral load according to time from infection regardless of age at study entry.

A major issue in modeling age and time in observational longitudinal data is selection or recruitment bias (16–19), that is, participants recruited at an older age may be different from participants who were recruited at a younger age and remained

in the study to become older persons. For example, participants who begin at age 60 may be systematically different from participants who started at age 40 and remain in the study for 20 years. For example, in the Baltimore Longitudinal Study of Aging (BLSA) (20), all volunteers are healthy at recruitment, whereas the health of participants who have been in the study for some time may have deteriorated over time. Repeated measures analyses must account for these differences.

In this article, we examine how different parameterizations of age and time in modeling observational longitudinal data can affect the results.

METHODS

Linear Mixed-effects Model

The linear mixed-effects (21) model for the vector of n_i observations for the i -th individual is

$$y_i = X_i\beta + Z_i b_i + \varepsilon_i,$$

where y_i is the $n_i \times 1$ vector of observations for individual i , and X_i and Z_i are the design matrices containing explanatory variables for the fixed (β) and random (b_i) effects. The fixed effects are regression parameters that estimate the average rate of change in the response variable for the differences in the fixed explanatory variables. The random effects account for the variability among individuals as well as the variability in the rates of change among individuals. It is assumed that the random errors $\varepsilon_i \sim N(0, \sigma^2 I)$, the random effects, $b_i \sim N(0, D)$, where D is a positive definite covariance matrix, and that ε_i and b_i are independent.

Study Population

The BLSA (20) is an ongoing open-panel prospective study of men and women conducted by the intramural research program of the National Institute on Aging with men participating since 1958 and women since 1978. Participants enter the study at various ages, have been in the study for varying lengths of time, have unequally spaced follow-up visits, and have an unequal number of measurements. Modeling is necessary to separate cross-sectional and longitudinal changes from this unbalanced data.

Linear mixed-effects models have been used in many analyses conducted on the BLSA population (22–29). In this study we consider models for 406 female and 202 male participants with an average of 2.8 repeated measurements (maximum of 6) for a total of 1,747 observations performed approximately every 2 years.

Jointly Modeling Cross-sectional and Longitudinal Change

Brant and colleagues have proposed models that estimate cross-sectional differences among participants along with the longitudinal changes within participants by considering for each individual i and examination time j , $\text{Age}_{ij} = \text{FAge}_i +$

Time_{ij} , where FAge_i , the age at first examination, is used to model cross-sectional age differences and Time_{ij} , the time since first examination, captures the longitudinal trend. In addition, interaction terms involving FAge_i and Time_{ij} are often necessary to adequately model a given data set.

Comparing and Testing Models for Cross-sectional and Longitudinal Change

Model 1: Simple age models with one random effect.—We first consider a simple model where the response y is modeled as a linear function of age with a single random effect for intercept. The simple age model is

$$y = \beta_0 + b_{i0} + \beta_1 \text{Age} + \varepsilon, \quad (1)$$

which can be rewritten as

$$\begin{aligned} y &= \beta_0 + b_{i0} + \beta_1 (\text{FAge} + \text{Time}) + \varepsilon \\ &= \beta_0 + b_{i0} + \beta_1 \text{FAge} + \beta_1 \text{Time} + \varepsilon. \end{aligned}$$

An alternative model, or first age and time model, with separate terms for first age and time is

$$y = \beta_0 + b_{i0} + \beta_1 \text{FAge} + \beta_2 \text{Time} + \varepsilon. \quad (2)$$

Note that Models 1 and 2 will be the same only if $\beta_1 = \beta_2$, and this hypothesis may be tested within the mixed-model framework. In the first age and time Model 2, the cross-sectional and longitudinal time effects are not constrained to be the same.

Model 2: Simple age models with two random effects.—Now suppose that the model also contains a random Age term in the age model or a random Time term in the first age and time model to allow participants to have varying longitudinal rates of change. In this case the age model is

$$y = \beta_0 + b_{i0} + (\beta_1 + b_{i1}) \text{Age} + \varepsilon, \quad (3)$$

which can be rewritten as

$$\begin{aligned} y &= \beta_0 + b_{i0} + (\beta_1 + b_{i1}) (\text{FAge} + \text{Time}) + \varepsilon \\ &= \beta_0 + b_{i0} + (\beta_1 + b_{i1}) \text{FAge} \\ &\quad + (\beta_1 + b_{i1}) \text{Time} + \varepsilon. \end{aligned}$$

The first age and time model is

$$y = \beta_0 + b_{i0} + \beta_1 \text{FAge} + (\beta_2 + b_{i1}) \text{Time} + \varepsilon. \quad (4)$$

Note that equation (4) allows the FAge and Time variables to have different fixed-effects coefficients and that the random coefficient is only included on the longitudinal Time term. Model 4 is not nested within Model 3 because in the age model the random slope term appears with both the FAge and Time terms when the model is expanded. Consequently, one cannot

use a Wald or likelihood ratio test to compare these two models but we can use either the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) (21,30–33). The model with the smaller AIC or BIC is preferred when using these criteria. It is important to estimate the mixed models using maximum likelihood, not restricted maximum likelihood (RML), because the models being compared do not contain the same fixed effects. RML adjusts the likelihood for the fixed effects in the model. Consequently, if the models do not contain the same fixed effects, RML will not produce comparable likelihood values. The best way to select among linear mixed-effects models based on various information criteria is still not clearly determined.

Model 3: Higher order models with one random effect.— In this model we allow a quadratic age effect with a single random effect for intercept. The model is

$$y = \beta_0 + b_{i0} + \beta_1 \text{Age} + \beta_2 \text{Age}^2 + \epsilon, \quad (5)$$

which can be rewritten as

$$\begin{aligned} y &= \beta_0 + b_{i0} + \beta_1 (\text{FAge} + \text{Time}) + \beta_2 (\text{FAge} + \text{Time})^2 + \epsilon \\ &= \beta_0 + b_{i0} + \beta_1 \text{FAge} + \beta_1 \text{Time} + \beta_2 \text{FAge}^2 \\ &\quad + 2\beta_2 \text{FAge} \times \text{Time} + \beta_2 \text{Time}^2 + \epsilon. \end{aligned}$$

The model involving first age and time is

$$\begin{aligned} y &= \beta_0 + b_{i0} + \beta_1 \text{FAge} + \beta_2 \text{Time} + \beta_3 \text{FAge}^2 \\ &\quad + \beta_4 \text{Time}^2 + \beta_5 \text{FAge} \times \text{Time} + \epsilon. \end{aligned} \quad (6)$$

To compare these models we need to test: $\beta_1 = \beta_2$ and $\beta_3 = \beta_4 = \frac{1}{2}\beta_5$.

Model 4: Higher order models with two random effects.— Finally, adding an Age or Time random effect to the higher order Model 5 yields

$$y = \beta_0 + b_{i0} + (\beta_1 + b_{i1}) \text{Age} + \beta_2 \text{Age}^2 + \epsilon, \quad (7)$$

which can be rewritten as

$$\begin{aligned} y &= \beta_0 + b_{i0} + (\beta_1 + b_{i1})(\text{FAge} + \text{Time}) \\ &\quad + \beta_2 (\text{FAge} + \text{Time})^2 + \epsilon \\ &= \beta_0 + b_{i0} + (\beta_1 + b_{i1}) \text{FAge} + (\beta_1 + b_{i1}) \text{Time} \\ &\quad + \beta_2 \text{FAge}^2 + 2\beta_2 \text{FAge} \times \text{Time} + \beta_2 \text{Time}^2 + \epsilon. \end{aligned}$$

The model involving first age and time is:

$$\begin{aligned} y &= \beta_0 + b_{i0} + \beta_1 \text{FAge} + (\beta_2 + b_{i1}) \text{Time} \\ &\quad + \beta_3 \text{FAge}^2 + \beta_4 \text{Time}^2 + \beta_5 \text{FAge} \times \text{Time} + \epsilon. \end{aligned} \quad (8)$$

These models are again not nested but may be compared using AIC or BIC.

Figure 1. Observed (thin lines) and modeled body weights of women with first age at least 70 years' old with no more than three repeated measurements. The thick dash-dot line is from the Age Model 1 and the thick solid lines are from the FAge and Time Model 2.

Note that all the models described above may contain additional explanatory variables without compromising the models (unless these additional variables have interaction terms with Age, FAge, or Time).

RESULTS

Model 1 Example: One Random Effect for Modeling Body Weight

In this example we wish to model the body weights of 43 women in the BLSA 70 years and older who had a short follow-up period (only two or three repeated measurements). The resulting data contain 103 observations with a mean of 2.4 repeated observations per participant (Figure 1). Note that we selected participants with at most three visits so that longitudinal patterns of change appear to be linear and hence equations (1) and (2) from Model 1 will best describe these data. Also, participants with at most three visits require a model with only a single random effect in order to provide a satisfactory description of the data. Appendix A provides the SAS code used to fit all the models. Table 1 compares the results of the fits of Model 1 to these data. The FAge and Time effects are statistically different ($p = .018$), and both the AIC and BIC suggest that the first age and time version of Model 1 (equation 2) is the most appropriate. The plot (Figure 1) of the fitted version of Model 1 (equations 1 and 2) clearly shows the differences between the model versions as well as the selection bias present in the data. Women beginning the study at older ages have lower predicted body weights than women who reached that age while already in the cohort. Interestingly, the Time coefficient is not statistically significant in the first age and time version ($p = .116$), suggesting that there are cross-sectional differences in body weight with age but no statistically significant longitudinal changes over such a short follow-up time.

Table 1. Results for Model 1 Applied to Weight Change Data in Women Aged 70 and Older

Age Model 1				FAge and Time Model 2			
Fixed Effects				Fixed Effects			
	Estimate	<i>t</i>	<i>p</i>		Estimate	<i>t</i>	<i>p</i>
Intercept	79.5327	9.69	<.001	Intercept	131.95	5.79	<.001
Age	-0.2418	-2.37	.021	FAge	-0.9285	-3.13	.003
				Time	-0.1698	-1.60	.116
				Test comparing models $H_0: \beta_F = \beta_T, t = -2.43 (p = .018)$			
Fit Statistics				Fit Statistics			
-2 lnL	654.1			-2 lnL	648.5		
AIC	662.1			AIC	658.5		
BIC	669.2			BIC	667.3		

Note: lnL = natural logarithm of the likelihood; AIC = Akaike information criterion; BIC = Bayesian information criterion.

The estimated random effects for the age Model 1 show a downward trend with Age ($p = .028$; figure not shown), suggesting the possible existence of unaccounted for terms missing from the model. In contrast, the estimated random effects for the first age and time Model 2 (figure not shown) are randomly scattered with no correlation ($p = 1.0$).

Model 2 Example: Two Random Effects for Modeling Cholesterol Change

This example considers cholesterol levels (mg/dL) of 202 male participants aged 60 and older who had between two and four visits resulting in 616 observations and an average of 3.0 observations per participant. Equations (3) and (4) are fit to these data. Table 2 gives a summary of the two models and Figure 2 plots the observed cholesterol levels as well as the fitted models. The plots (Figure 2) indicate that the longitudinal and cross-sectional changes are different, probably because of selection bias due to the recruitment of healthy volunteers. Again both the AIC and BIC prefer the first age and time model Version 4 to the age model Version 3. In the first age and time model, the FAge term is not statistically significant ($p = .071$), suggesting no statistically significant cross-sectional differences but longitudinal declines in cholesterol.

Model 3 Example: Higher Order Terms and One Random Effect for Modeling Body Weight

Often one has data in which only a limited number of repeated measurements exist for each person and there are

also definite age/time effects. Consider modeling the weight of the 198 BLSA women volunteers who had exactly two visits, thus making the inclusion of a random Age or Time term unnecessary. Here, we include a random intercept as well as the fixed components given in Model 3 with an additional Age² term in Model 5 or additional FAge², Time², and FAge $\times$ Time terms in Model 6. Although there is not much follow-up time and it is not anticipated that the fixed-effects Time² term will be significant, it is retained in the model to make the first age and time model comparable with the age model. The results of the analysis are provided in Table 3 as well as in Figure 3. As expected, the Time² coefficient in Model 6 is not statistically significant ($p = .304$); but the FAge² and FAge $\times$ Time coefficients are significant ($p < .001$ and $p = .007$, respectively), and the Age² coefficient in Model 5 is statistically significant ($p < .001$). The test of the contrast of the parameters ($p < .001$) and both the AIC and BIC clearly indicate a preference for the first age and time model over the age model.

Model 4 Example: Higher Order Terms and Two Random Effects for Modeling Maximum Systolic Blood Pressure Achieved During Exercise

A treadmill test is performed in the BLSA on participants without severe contraindications (26). One variable measured during this test is the maximum systolic blood pressure (MaxSBP). We use data from 274 women aged 40 and

Table 2. Results for Model 2 Applied to Cholesterol Change Data in Men Aged 60 and Older

Age Model 3				FAge and Time Model 4			
Fixed Effects				Fixed Effects			
	Estimate	<i>t</i>	<i>p</i>		Estimate	<i>t</i>	<i>p</i>
Intercept	327.96	14.95	<.001	Intercept	266.53	10.63	<.001
Age	-1.4814	-5.22	<.001	FAge	-0.6150	-1.81	.071
				Time	-2.6613	-6.53	<.001
Fit statistics				Fit statistics			
-2 lnL	5971.2			-2 lnL	5957.5		
AIC	5983.2			AIC	5971.5		
BIC	6003.0			BIC	5994.6		

Note: lnL = natural logarithm of the likelihood; AIC = Akaike information criterion; BIC = Bayesian information criterion.

Figure 2. Observed (thin lines) and modeled cholesterol levels (mg/dL) of men with first age at least 60 years' old with 2 to four repeated measurements. The thick dash-dot line is from the Age Model 3 and the thick solid lines are from the FAge and Time Model 4.

older who have at most five measurements of MaxSBP. The data set contains 744 observations resulting in 2.72 observations per participant. When equations (7) and (8) were fitted in Model 4 (see Table 4) the fitted lines for both equations (Figure 4) are almost indistinguishable, suggesting no selection bias. The AIC prefers the first age and time model, whereas the BIC prefers the age model. In the first age and time model, the FAge² term is not statistically significant ($p = .069$) and could be eliminated from the model. The estimated random effects are not significantly associated with Age for either model ($p = 1.0$, $p = 1.0$, $p = .808$, and $p = .915$), suggesting that, in this case, either model provides a good description of the data (figures not shown).

DISCUSSION

In studying longitudinal data, many researchers find it favorable to use a single term for age in their statistical models, thus exhibiting the effect of age as a continuous curve

Figure 3. Observed (thin lines) and modeled body weights of women with two repeated measurements. The thick dash-dot line is from the Age Model 5 and the thick solid lines are from the FAge and Time Model 6.

(eg, see thick broken lines in Figures 1–4) However, this approach makes no distinction between cross-sectional and longitudinal differences. Often these same researchers find objectionable or question a graph such as the thick solid lines in Figures 1–3 where it appears that persons of the same age have different estimated response values. This approach fails to consider the possibility of a recruitment bias effect which may occur by design (healthy volunteers) or because sick individuals are less likely to be enrolled in a longitudinal study. For example, a healthy person who starts the study at age 40 when evaluated 10 years later may have less favorable risk values than a healthy 50-year-old individual of the same sex. This problem is more likely to be a concern in observational longitudinal studies that recruit individuals over a broad age range and perform relatively long follow-up.

In this article, data from observational longitudinal studies are analyzed using linear mixed-effects regression models with terms involving entry age and time. Other approaches for analyzing longitudinal data can be used depending on

Table 3. Results for Model 3 Applied to Weight Change Data in Women Aged 70 and Older

Age Model 5				FAge and Time Model 6			
Fixed Effects				Fixed Effects			
	Estimate	<i>t</i>	<i>p</i>		Estimate	<i>t</i>	<i>p</i>
Intercept	35.8896	6.89	<.001	Intercept	39.2132	5.38	<.001
Age	1.1570	5.71	<.001	FAge	1.1565	3.84	<.001
Age ²	−0.0098	−5.09	<.001	Time	1.2394	3.96	.001
				FAge ²	−0.01127	−3.84	<.001
				Time ²	−0.02181	−1.03	.304
				FAge × Time	−0.01329	−2.71	.007
				Test comparing models: $H_0: \beta_F = \beta_T$ & $\beta_{F \times T} = \beta_{T \times T} = \frac{1}{2}\beta_{F \times T}$			
				$F = 8.03$ ($p < .001$)			
Fit statistics				Fit statistics			
−2 lnL	2819.0			−2 lnL	2795.7		
AIC	2829.0			AIC	2811.7		
BIC	2845.5			BIC	2838.0		

Note: lnL = natural logarithm of the likelihood; AIC = Akaike information criterion; BIC = Bayesian information criterion.

Table 4. Results for Model 4 Applied to Maximum Systolic Blood Pressure Data During Exercise

Age Model 7				FAge and Time Model 8			
Fixed Effects				Fixed Effects			
	Estimate	t	p		Estimate	t	p
Intercept	54.4727	2.44	.015	Intercept	57.2684	1.66	.098
Age	3.1002	4.31	<.001	FAge	3.0133	2.60	.010
Age ²	−0.01780	−3.12	.002	Time	3.6822	3.95	<.001
				FAge ²	−0.01734	−1.83	.069
				Time ²	−0.05472	−2.19	.029
				FAge × Time	−0.03738	−2.53	.012
Fit statistics				Fit statistics			
−2 lnL	6570.6			−2 lnL	6559.7		
AIC	6584.6			AIC	6579.7		
BIC	6609.9			BIC	6615.8		

Note: lnL = natural logarithm of the likelihood; AIC = Akaike information criterion; BIC = Bayesian information criterion.

the specific question to be addressed. Hierarchical or multi-level models (6,7,34) are used to describe time trends in a well-defined population with specified, fixed-time intervals; logistic regression may be used to model the probability of some event occurring in a fixed follow-up time period as a function of explanatory variables; and the Cox proportional hazards model is used when time to event is of interest (34). In addition, other authors have recognized the importance and flexibility of random effects or mixed-effects models in analyzing longitudinal data. Gueorguieva and Krystal (35) describe how mixed-effects models have become the preferred choice of analysis for repeated measures data in psychiatric research, whereas Morrell et al. (36) discuss the importance of variable selection that maintains a hierarchy in the fixed and random terms in the mixed-effects model.

In this article, we have examined the implications of modeling the longitudinal and cross-sectional components in a linear mixed-effects model using either age or first age and time terms. The examples presented show that for observational longitudinal studies covering a broad age span, the

first age and time approach is more flexible and in many cases will provide a better description of the data than using only age terms.

APPENDIX A: SAS CODE USED TO FIT THE MODELS

Model 1 Example.

```
proc mixed method=ml noclprint;
class idno;
model wgt= fage time / solution;
random int / subject=idno type=un g;
estimate 'fage=time' fage 1 time -1;
contrast 'fage=time' fage 1 time -1;
title 'FAge and Time Model for Model 1 Example';
run;
proc mixed method=ml noclprint;
class idno;
model wgt= age / solution;
random int / subject=idno type=un g;
title 'Age Model for Model 1 Example';
run;
```

Model 2 Example.

```
proc mixed method=ml noclprint;
class idno;
model chol= fage time / solution;
random int time / subject=idno type=un g;
title 'FAge and Time Model for Model 2 Example';
run;
proc mixed method=ml noclprint;
class idno;
model chol = age / solution;
random int age / subject=idno type=un g;
title 'Age Model for Model 2 Example';
run;
```

Model 3 Example.

```
proc mixed method=ml noclprint;
class idno;
model wgt= fage time fage*fage time*time fage*time /
solution;
```

Figure 4. Observed (thin lines) and modeled maximum systolic blood pressure (mmHg) of women aged 40 and older with at most five repeated measurements. The thick dash-dot line is from the Age Model 7 and the thick solid lines are from the FAge and Time Model 8.

```

random int / subject=idno type=un g;
contrast 'quadratic' fage 1 time -1,
fage*fage 1 time*time -1 ,
fage*time 0.5 time*time -1 ;
title 'FAge and Time Model for Model 3 Example ';
run;
proc mixed method=ml noclprint;
class idno;
model wgt= age*age / solution;
random int / subject=idno type=un g;
title 'Age Model for Model 3 Example ';
run;
Model 4 Example.
proc mixed method=ml noclprint;
class idno;
model maxsbp= fage time fage*fage time*time fage*time
/ solution;
random int time / subject=idno type=un g;
title 'FAge and Time Model for Model 4 Example ';
run;
proc mixed method=ml noclprint;
class idno;
model maxsbp= age*age / solution;
random int age / subject=idno type=un g;
title 'Age Model for Model 4 Example ';
run;

```

FUNDING

This research was supported by the Intramural Research Program of the National Institutes of Health, National Institute on Aging.

ACKNOWLEDGMENTS

The authors thank the reviewers for their comments that led to a substantial improvement in the article.

CORRESPONDENCE

Address correspondence to Larry J. Brant, PhD, Gerontology Research Center, National Institute on Aging, 5600 Nathan Shock Drive, Baltimore, MD 21224. Email: larry_brant@nih.gov

REFERENCES

1. Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. *Am J Epidemiol.* 2005;161:280–288.
2. Brant LJ, Verbeke GN. Describing the natural heterogeneity of aging using multilevel regression models. *Int J Sports Med.* 1997;18:S225–S231.
3. Schaie KW. A general model for the study of developmental problems. *Psychol Bull.* 1965;64:92–107.
4. Lindstrom MJ, Bates DM. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *J Am Stat Assoc.* 1988;83:1014–1022.
5. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics.* 1982;38:963–974.
6. Bryk AS, Raudenbush SW. Application of hierarchical linear models to assessing change. *Psychol Bull.* 1987;101:147–158.
7. Goldstein H. Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika.* 1986;73:43–56.
8. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. *Applied Regression and Other Multivariable Methods.* Pacific Grove, CA: Duxbury; 1998.
9. Little RJA, Rubin DB. 1987. *Statistical Analysis With Missing Data.* New York, NY: John Wiley.
10. Alfaro-Acha A, Snih SA, Raji MA, et al. Handgrip strength and cognitive decline in older Mexican Americans. *J Gerontol A Biol Sci Med Sci.* 2006;61:859–865.
11. Miller RR, Cappola AR, Shardell MD, et al. Persistent changes in interleukin-6 and lower extremity function following hip fracture. *J Gerontol A Biol Sci Med Sci.* 2006;61:1053–1058.
12. Miller RR, Hicks GE, Shardell MD, et al. Association of serum vitamin D levels with inflammatory response following hip fracture: the Baltimore hip studies. *J Gerontol A Biol Sci Med Sci.* 2007;62:1402–1406.
13. Blaum CS, West NA, Haan MN. Is the metabolic syndrome, with or without diabetes, associated with progressive disability in older Mexican Americans? *J Gerontol A Biol Sci Med Sci.* 2007;62:766–773.
14. Park T, Lee SY. Simple pattern-mixture models for longitudinal data with missing observations: analysis of urinary incontinence data. *Stat Med.* 1999;18:2933–2941.
15. Edwards LJ, Stewart PW, MacDougall JE, Helms RW. A method for fitting regression splines with varying polynomial order in the linear mixed model. *Stat Med.* 2006;25:513–527.
16. Heckman JJ, Ichimura H, Smith J, Todd P. Sources of selection bias in evaluating social programs: an interpretation of conventional measures and evidence on the effectiveness of matching as a program evaluation method. *Proc Natl Acad Sci U S A.* 1996;93:13416–13420.
17. Jensen TK, Scheike T, Keiding N, Schaumburg I, Grandjean P. Selection bias in determining the age dependence of waiting time to pregnancy. *Am J Epidemiol.* 2000;152:565–572.
18. Korn EL, Teeter DM, Baumrind S. Using explicit clinician preferences in nonrandomized study designs. *J Stat Plann Inference.* 2001;96:67–82.
19. Hogan JW, Lin X, Herman B. Mixtures of varying coefficient models for longitudinal data with discrete or continuous nonignorable dropout. *Biometrics.* 2004;60:854–864.
20. Shock NW, Greulich RC, Andres R, et al. 1984. *Normal Human Aging: The Baltimore Longitudinal Study of Aging (NIH Publication No. 84–2450).* Washington, DC: U.S. Government Printing Office.
21. Verbeke G, Molenberghs G. 2000. *Linear Mixed Models for Longitudinal Data.* New York, NY: Springer-Verlag.
22. Morrell CH, Brant LJ. Modeling hearing thresholds in the elderly. *Stat Med.* 1991;10:1453–1464.
23. Morrell CH, Pearson JD, Brant LJ, Gordon-Salant S. Construction of hearing percentiles in women with non-constant variance from the linear mixed-effects model. *Stat Med.* 1997;16:2475–2488.
24. Pearson JD, Morrell CH, Landis PK, Carter HB, Brant LJ. Mixed-effects regression models for studying the natural history of prostate disease. *Stat Med.* 1994;13:587–601.
25. Brant LJ, Sheng SL, Morrell CH, Verbeke GN, Lesaffre E, Carter HB. Screening for prostate cancer by using random-effects models. *J R Stat Soc A.* 2003;166:51–62.
26. Fleg JL, Morrell CH, Bos AG, et al. Accelerated longitudinal decline of aerobic capacity in healthy older adults. *Circulation.* 2005;112:674–682.
27. Pearson JD, Morrell CH, Brant LJ, Landis PK, Fleg JL. Age-associated changes in blood pressure in a longitudinal study of healthy men and women. *J Gerontol A Biol Sci Med Sci.* 1997;52:M177–M183.
28. Brant LJ, Sheng SL, Morrell CH, Zonderman AB. Data from a longitudinal study provided measurements of cognition to screen for Alzheimer's disease. *J Clin Epidemiol.* 2005;58:701–707.
29. Morrell CH, Brant LJ, Pearson JD, Verbeke GNM, Fleg JL. Applying linear mixed-effects models to the problem of measurement error in epidemiologic studies. *Comm Statist Simul Comput.* 2003;32:437–459.
30. Dziak JJ, Li R. 2007. An overview on variable selection for longitudinal data. In: Hong D, ed. *Quantitative Medical Data Analysis.* Singapore: World Sciences; 2007.

31. Gomez EV, Schaalje GB, Fellingham GW. Performance of the Kenward–Roger method when the covariance structure is selected using AIC and BIC. *Comm Statist Simul Comput*. 2005;34:377–392.
32. Gurka MJ. Selecting the best linear mixed model under REML. *Am Stat*. 2006;60:19–26.
33. Satkartar KK, Dunson DB. Fixed and random effects selection in linear and logistic models. *Biometrics*. 2007;63:690–698.
34. Singer JD, Willett JB. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York, NY: Oxford University Press.
35. Gueorguieva R, 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.
36. Morrell CH, Pearson JD, Brant LJ. Linear transformations of linear mixed-effects models. *Am Stat*. 1997;51:338–343.

Received March 5, 2008

Accepted June 26, 2008

Decision Editor: Darryl Wieland, PhD, MPH