Introduction to Longitudinal Data Analysis Using the Traumatic Brain Injury Model Systems National Database

Prepared by Jessica M. Ketchum from the Traumatic Brain Injury Model Systems National Data and Statistical Center

Methodology described in these Training Courses is primarily taken from two sources, with additional references cited throughout:

- Chapters 4 and 5 in Hedeker and Gibbons (2006), Longitudinal Data Analysis, Wiley
- Hedeker (2004) Chapter 12: An Introduction to growth modeling, In D. Kaplan (Ed.) Quantitative Methodology for Social Sciences. Thousand Oaks CA: Sage.

Dr. Hedeker and his colleagues have done an excellent job describing methodology for longitudinal data analysis (LDA) to researchers in the biomedical and behavioral science field with the least amount of technicality needed while emphasizing critical concepts with applications. As such, there is no need to reword much of what he has so eloquently described. With the Traumatic Brain Injury Model Systems (TBIMS) researchers in mind, both with novice and advanced statistical training, I have expanded on some parts of his methodical description, condensed others, and applied the methodology to a large sample of subjects in the TBIMS National Database (NDB).

These Training Courses are presented as a series and have been developed to educate TBIMS researchers on the different concepts of LDA starting with the most basic extension from simple linear regression and working towards more advanced concepts. A webinar is being prepared for each Training Course and will be presented to TBIMS researchers through the TBIMS Analytic Special Interest Group (SIG). All presentations, course notes, data files, and SAS code will be available for download through the TBIMS NDSC website (https://www.tbindsc.org/Researchers.aspx). Questions on course material or individual consultation for research projects involving LDA and the TBIMS NDB should be directed to Dr. Ketchum (jketchum@craighospital.org). While we do our best to respond to all questions and requests, priority is given to TBIMS internal researchers (including Committees, Modules, and SIGs), followed by external researchers with submitted requests for use of the TBIMS NDB, and last by external researchers looking to apply these methods to future projects with the TBIMS NDB.

LDA Training Course 1: Extending Simple Linear Regression Model to Mixed-Effects Regression Model

The Traumatic Brain Injury Model Systems (TBIMS) National Database (NDB) is the largest longitudinal study on TBI outcomes in the world. Perhaps the most salient characteristic of the TBIMS NDB is its longitudinal nature, containing valuable information regarding both between and within subject change in outcomes over time. One of the most useful strategies for analyzing these data are methods for longitudinal data analysis (LDA). A unique characteristic of LDA is its ability to account for the correlations in repeated measures of subjects over time. This

training course as a whole aims to guide the researcher through the different concepts of LDA, starting with the most basic extension of LDA from simple linear regression. We will build upon this basic model using the TBIMS NDB for demonstration. As the goal is to build upon our knowledge, examples using the same outcome will be carried through as extensions are made; alternative outcomes and covariates will be introduced as needed in later Training Courses.

Introduction

Longitudinal studies are very common in social science research and the methods to analyze longitudinal data are seeing increasing use in rehabilitation research, particularly using the TBIMS NDB. In longitudinal studies, subjects are measured repeatedly over time, and interest is often focused on characterizing their change, or growth, across time.

Traditional analysis of variance (ANOVA) methods for growth curve analyses are described in Bock (1975); however, these methods are of limited use due to restrictive assumptions concerning missing data across time and the variance-covariance structure of the repeated measures. The univariate mixed-model analysis of variance assumes that the variances and covariances of the dependent variable across time are equal (i.e., compound symmetry). The multivariate analysis of variance for repeated measures only includes subjects with complete data across all time points. Also, these traditional methods primarily focus on estimation of group trends across time and provide little help in understanding how specific individuals change over time. For these reasons, **mixed-effects regression models** (MRMs) have become the methods of choice for modeling longitudinal data.

Variants for MRMs have been developed under a variety of names:

- random-effects models (Laird & Ware, 1982);
- variance component models (Dempster, Rubin, & Tsutakaw, 1981);
- multilevel models (Goldstein, 1995);
- hierarchical linear models (Bryk & Raudenbush, 1992);
- two-stage models (Bock, 1989a);
- random coefficient models (de Leeuw & Kreft, 1986);
- individual growth curves (Bock & Thissen, 1980, Goldstein, 1981);
- mixed models (Longford, 1987; Wolfinger, 1993);
- empirical Bayes models (Hui & Berger, 1983; Strenio, Weisberg, &Bryk, 1983); and
- random regression models (Bock, 1983a; 1983b; Gibbons, Hedeker, Waternaux, & Davis, 1988).

A basic characteristic of all of these models is the inclusion of random subject effects into the regression model framework in order to account for the influence of subjects on their repeated observations. These random-effects describe each person's trend across time and explain the correlational structure of the longitudinal data. Additionally, they indicate the degree of subject variation that exists in the population of subjects.

Several features make MRMs especially useful in longitudinal research. First, subjects are not assumed to be measured on the same number of time points; thus, subjects with incomplete data across time are included in the analysis. This is an important advantage relative to

procedures that require complete data across time for two reasons: (a) by including all data the analysis has increased statistical power, and (b) complete case analysis may suffer from biases to the extent that subjects with complete data are not representative of the larger population of subjects. A second important feature is that time is modeled as a continuous variable, so subjects do not need to be measured at the same time points. This is useful in longitudinal studies when follow-up times are not uniform across subjects. Third, both time-invariant and time-variant covariates can be included in the model. Thus, changes in the outcome variable may be due to both stable characteristics with time (e.g., sex or race) as well as characteristics that change across time (e.g., changing functional status, life events). Finally, whereas traditional approaches estimate average change across time in a population, MRM can also estimate change for each subject. These estimates of individual change across time can be particularly useful in longitudinal studies, where a proportion of subjects exhibit change across time that deviates from the average trend.

These LDA Training Courses will focus on describing MRMs for continuous outcomes in a very practical way. We will illustrate how MRMs can be seen as an extension of an ordinary (simple) linear regression model. Starting here, the model will be slowly extended and described, guiding the reader from more familiar to less familiar territory. Following a description of the statistical model and extension presented, two MRM example analyses will be presented using the TBIMS NDB. As we further develop the model to handle more complexities, these example analyses will illustrate many of the key features of MRMs for longitudinal data analysis that can be applied using the TBIMS NDB and other longitudinal research studies.

A Simple Linear Regression Model

To introduce MRMs, consider a simple linear regression model for the measurement *y* of individual *i* (*i* = 1, 2, …, *N* subjects) on occasion *j* (*j* = 1, 2, …, n_i occasions):

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \varepsilon_{ij}. \quad (1)$$

Ignoring subjects, this model represents the regression of the outcome variable *y* on the independent variable time (denoted by *t*). The subscripts keep track of the particulars of the data, namely whose observation it is (subscript *i*) and when this observation was made (subscript *j*). The independent variable *t* gives a value to the level of time, and may represent time in days, weeks, months, years, etc. Since *y* and *t* carry both *i* and *j* subscripts, both the outcome variable and the time variable are allowed to vary by individuals and occasions.

In linear regression models, the errors ε_{ij} are assumed to be normally and *independently* distributed in the population with mean of zero and variance σ^2. The independence assumption makes the model given in (1) an unreasonable one for longitudinal data. This is because the outcomes *y* are observed repeatedly from the same individual, and so it is much more reasonable to assume that the errors within an individual are correlated to some degree. Furthermore, the above model assumes that the growth, or change across time, is the same for all individuals because the model parameters describing growth (β_0, the intercept or initial level, and β_1, the linear change across time) do not vary by individuals. For both of these reasons, it is useful to add individual-specific effects into the model that will account for the data dependency and describe differential growth for different individuals. This is precisely

what MRMs accomplish. The essential point is that MRMs can be viewed as augmented linear regressions models.

Random-Intercept Mixed Regression Model

A simple extension of the regression model given in (1) to allow for the influence of each individual on their repeated outcome is provided by

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \upsilon_{0i} + \varepsilon_{ij}, \quad (2)$$

where υ_{0i} represents the influence of individual *i* on their repeated observations. Notice that if individuals have no influence on their repeated outcomes, then all of the υ_{0i} terms would equal 0. However, it is more likely that repeated observations for a given subject will be correlated with either a positive or negative association, and so the υ_{0i} terms will deviate from 0.

To better reflect how this model characterizes an individual's influence on their observations, it is helpful to represent the model in a hierarchical linear model (HLM) or multilevel form [Goldstein, 1995; Raudenbush and Bryk, 2002]. For this, it is partitioned into the within-subjects (or Level 1) model,

$$y_{ij} = b_{0i} + b_{1i} t_{ij} + \varepsilon_{ij}, \quad (3)$$

and the between-subjects (or Level 2) model,

$$b_{0i} = \beta_0 + \upsilon_{0i}, \quad (4)$$
$$b_{1i} = \beta_1.$$

Here, the Level 1 model indicates that individual *i*'s response at time *j* is influenced by their initial level b_{0i} and time trend, or slope, b_{1i}. The Level 2 model indicates that the individual *i*'s initial level (intercept) is determined by the population initial level β_0, plus a unique contribution for that individual υ_{0i}. Thus, each individual has their own distinct level. Conversely, the present model indicates that each individual's slope is the same; all are equal to the population slope β_1. Another way to think about this is that each person's trend line is parallel to the population trend determined by β_0 and β_1. The difference between each individual's trend and the population trend is υ_{0i}, which is constant across time.

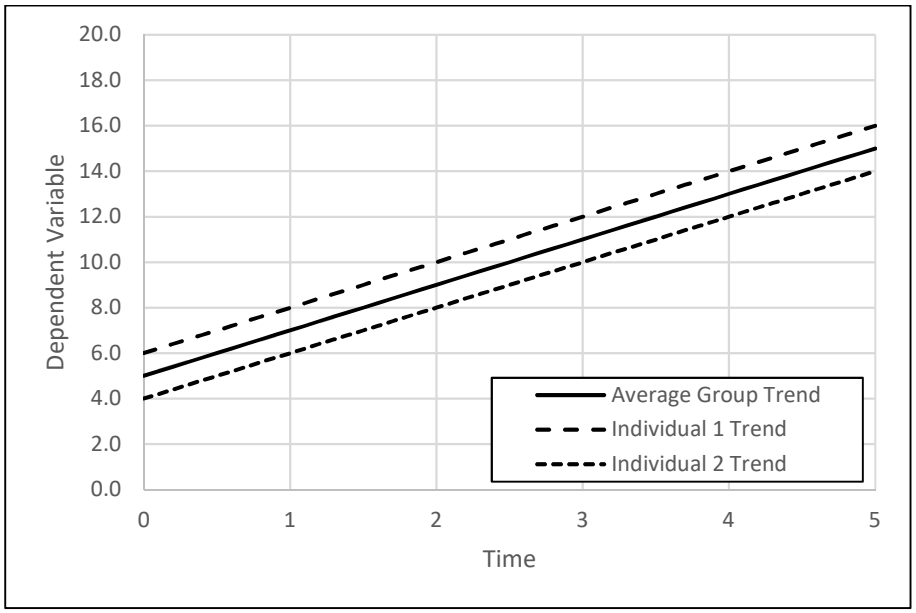
The between-subjects, or Level 2, model is sometimes referred to as a "slopes as outcomes" model (Burstein, 1980). The HLM representation shows that just as within-subjects (Level 1) covariates can be included in the model to explain variation in Level 1 outcomes (y_{ij}), between-subjects (Level 2) covariates can also be included to explain variation in Level 2 outcomes (the subject's intercept b_{0i} and slope b_{1i}). Note that combining the within- and between-subject models (3) and (4) yields the previous single-equation model (2).

Because individuals in a sample are typically thought to be representative of a larger population of individuals, the individual-specific effects υ_{0i} are treated as "random-effects". That is, υ_{0i} are considered to be representative of a distribution of individuals in in the population. The most common form for this population distribution is the normal distribution, with mean 0 and variance σ_υ^2. In the model given by equation (2), the errors ε_{ij} are now assumed to be *conditionally independently* distributed in the population with zero mean and common variance σ^2. Conditional independence here means conditional on the random individual-specific effects

υ_{0i}. Because the errors now have an influence due to the individuals removed from them, this conditional independence assumption is much more reasonable than the ordinary independence assumption associated with the simple linear regression model in (1).

A graphical representation of the random intercept MRM is shown below in Figure 1.

Figure 1: Random Intercept Model



Here, individuals deviate from the regression of *y* on *t* in a parallel manner since there is only one subject effect υ_{0i}. Thus, it is often referred to as a "random-intercepts model", with each υ_{0i} indicating how individual *i* deviates from the population trend. In this figure, the solid line represents the population average trend, which is based on β_0 and β_1. Also depicted are two individual trends, one below and one above the population trend, shown with different dashed lines. For a given sample, there is a line for each individual in the sample. The variance term σ_υ^2 represents the spread of these lines from the population line. If σ_υ^2 is near zero, then the individual lines would not deviate much from the population trend. In that case, individuals do not exhibit much heterogeneity in their change across time. Alternatively as individuals differ from the population trend, the lines move away from the population trend line and σ_υ^2 increases. In this case, there is more individual heterogeneity in time trends.

Compound Symmetry and Intraclass Correlation

The random-intercept model implies a *compound symmetric* association for the variances and the covariances of the longitudinal data. That is, the variances and covariances are assumed to be the same, namely

$$V(y_{ij}) = \sigma_\upsilon^2 + \sigma^2, \tag{5}$$

$$Cov(y_{ij}, y_{ij'}) = \sigma_\upsilon^2, \qquad \text{where } j \neq j'$$

Thus, the variance of any observation is expressed as $\sigma_\upsilon^2 + \sigma^2$, and the covariance between any two observations from the same subject at different timepoints is expressed as σ_υ^2. As subjects

are assumed to be independent, the covariance between observations from different subjects, irrespective of time, will be 0.

The compound symmetry pattern for the variance-covariance in matrix notation for 4 time points for subject *i* is expressed as:

$$\mathbf{V}_i = \begin{bmatrix} \sigma_v^2 + \sigma^2 & \sigma_v^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma^2 & \sigma_v^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 & \sigma_v^2 + \sigma^2 & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 & \sigma_v^2 & \sigma_v^2 + \sigma^2 \end{bmatrix}.$$

Expressing the covariance as a correlation yields the **intraclass correlation (ICC)**, which is the ratio of the individual variance σ_v^2 to the total variance $\sigma_v^2 + \sigma^2$ [i.e., $ICC = \sigma_v^2/(\sigma_v^2 + \sigma^2)$]. This coefficient represents the degree of association of the longitudinal data within subjects, and specifically indicates the proportion of variance in the data attributable to individuals.

Inference

Hypothesis testing for the fixed effects parameters (i.e., β) generally involve the so-called "Wald test" [Wald, 1943], which uses the ratio of the parameter estimate to its standard error to determine statistical significance. These tests statistics (i.e., *Z* = ratio of the parameter estimate to its standard error) are compared to a standard normal frequency table to test the null hypothesis that the parameter is equal to 0. Alternatively, these *Z*-statistics are sometimes squared, in which case the resulting test statistic is distributed as chi-square on one degree of freedom. In either case, the *p*-values are identical.

For the variance and covariance terms, there are concerns in using the standard errors in constructing Wald test statistics, particularly when the population variance is thought to be near zero and the number of subjects is small [Bryk and Raudenbush, 1992]. This is because variance parameters are bounded; they cannot be less than zero and so using standard normal for the sampling distribution is not reasonable. As a result, we will not include the Wald tests for variance and covariance terms in these Training Courses.

When comparing nested models, the likelihood ratio test can be used to perform uni- or multi-parameter hypothesis tests. For this, one compares the difference in model deviance values (i.e., -2 log L) to a chi-square distribution, where the degrees of freedom equal the difference in the number of parameters between the two models. It should be noted that while use of the likelihood ratio tests for fixed effects is not problematic, for variance and covariance terms this test also suffers from the variance boundary problems mentioned above [Verbeke and Molenberghs, 2000]. Based on simulation studies, it can be shown that the likelihood ratio test is too conservative (for testing null hypotheses about variance and covariance parameters), namely, it does not reject the null hypothesis often enough. This would then lead to accepting a more restrictive variance-covariance structure than is correct. As noted by Berkhof and Snijders [2001], this bias can be largely corrected by dividing the *p*-value obtained from the likelihood ratio test (of variance and covariance parameters) by two.

TBIMS National Dataset

Throughout these LDA Training Courses we will consider data from the Traumatic Brain Injury Model Systems (TBIMS) National Database (NDB). A detailed description of the TBIMS longitudinal study can be found at https://www.tbindsc.org/. Briefly, consenting participants who have had a moderate to severe TBI are enrolled in the TBIMS NDB during their inpatient rehabilitation stay at one of 23 (16 currently funded) TBIMS Centers in the United States. During their inpatient rehabilitation treatment stay, pre-injury, injury, and rehabilitation characteristics are collected by trained data collectors through medical record abstraction and patient/family interview. Follow-up interviews assessing functional, social, emotional, and medical outcomes are conducted by trained data collectors following standardized procedures at 1, 2, 5 years post-injury, and every 5 years thereafter. Currently (April 2021), the TBIMS NDB has data from over 18,500 participants and more than 70,000 follow-up interviews, with the longest follow-up out to 30 years post-injury, making it the largest longitudinal database on moderate to severe traumatic brain injury in the world. In this LDA Training Course, we will focus on the longitudinal relationship of Satisfaction with Life Scale (SWLS) over time 1 to 10 years post injury. SWLSTOT is the total score derived by summing the 5 SWLS items (each scored 1-7) and ranges from 5 to 35, with higher scores indicative of higher satisfaction with life.

A de-identified, limited analytic dataset has been prepared in SAS and can be downloaded from https://www.tbindsc.org/Researchers.aspx. In this dataset, SWLSTOT is self-report from participants (not proxy interview) at follow-up years 1, 2, 5, and 10 post-injury. There is a 2 month window around year 1 follow-up, a 3 month window around year 2 follow-up, and a 6 month window around 5 and 10 year follow-ups. Specific dates for injury and follow-up are available and could be used to compute time post-injury more specifically for each subject; however, for simplicity TIME will be considered to be 1, 2, 5, and 10 years for each subject. We selected participants with injury dates between March 1996 and September 2007, so that 10 year follow-up data on SWLS would be due for all subjects in the sample. We further selected participants who were followed (not by proxy) and had complete SWLS data for at least 2 (of 4) time points. The analytic data set includes a total of 4130 individuals. Age at injury (AGE), sex from medical records (SEX), and self-reported race/ethnicity (RACE) were collected during inpatient rehabilitation and will be used as time invariant covariates. These variables are summarized in Figure 2. Additional variables in the dataset will be introduced later in the Training Courses as needed.

Figure 2: Summary of TBIMS Analytic Data Set

TBIMS LDA Analytic Dataset

CenterID = values range from 1 to 23 (de-identified; not the same as TBIMS CenterID)

SubjectID = 1, 2, …, N_k (where N_k represents the total number of subjects for the k^{th} center; de-identified and recoded so that SubjectID is nested within CenterID)

CSubID = Center || SubID

FirstObs = 1 if first row for subject; 0 otherwise

SWLSTOT = SWLS score at follow-up (5-35); continuous

TIME = 1, 2, 5, 10; continuous

AGE = Age at injury (range 16-99); continuous

SEX = Male (0), Female (1); dichotomous

RACE = White (0), Black (1), Hispanic (2), Other (3); categorical

TIMEY1CENT = TIME – 1 = 0, 1, 4, 9 (so that time = 0 represents Year 1); continuous

AGECENT = AGE – 34.78 (sample mean age); range -18.78 to 52.22; continuous

The analytic dataset is presented in a "stacked" format, with each row representing a unique time point for each subject. CenterID has been deidentified and ranges from 1-23. We also recoded SubjectID to be nested within CenterID. That is, SubjectID = 1, 2, 3, …., N_k for all centers. This is done to de-identify the data but also to improve efficiency of model estimation in SAS. A subset of the data for the first two subjects from the first two centers is shown below in Figure 3.

Figure 3: Sample of Analytic Data

CenterID	SubID	CSubID	FirstObs	Time	SWLSTot	Age	Sex_new	Race_New	TimeY1Cent	AgeCent
1	2	1 2	1	1	15	67	0	0	0	32.2184
1	2	1 2	0	2	.	67	0	0	1	32.2184
1	2	1 2	0	5	20	67	0	0	4	32.2184
1	2	1 2	0	10	.	67	0	0	9	32.2184
1	3	1 3	1	1	22	54	0	0	0	19.2184
1	3	1 3	0	2	23	54	0	0	1	19.2184
1	3	1 3	0	5	24	54	0	0	4	19.2184
1	3	1 3	0	10	35	54	0	0	9	19.2184
2	1	2 1	1	1	23	19	0	0	0	-15.7816
2	1	2 1	0	2	9	19	0	0	1	-15.7816
2	1	2 1	0	5	23	19	0	0	4	-15.7816
2	1	2 1	0	10	22	19	0	0	9	-15.7816
2	2	2 2	1	1	35	47	0	3	0	12.2184
2	2	2 2	0	2	34	47	0	3	1	12.2184
2	2	2 2	0	5	30	47	0	3	4	12.2184
2	2	2 2	0	10	31	47	0	3	9	12.2184

In Table 1, we present the observed sample sizes, SWLS means, standard deviations, and percentiles (0, 25, 50, 75, 100) across the 4 study time points for the sample of 4130 subjects. From these data, we see the mean SWLS increases slightly over time (21.46 to 22.39). Additionally, it appears from the standard deviations and percentiles that the spread remains relatively stable over time (8.21-8.31).

Table 1: Observed SWLS Means, Standard Deviations (SD), and N across Time

Time	N	Mean SWLS (SD)	SWLS Percentiles
Year 1	3569	20.90 (8.28)	[5, 14, 22, 28, 35]
Year 2	3531	21.55 (8.34)	[5, 14, 23, 29, 35]
Year 5	3329	22.03 (8.34)	[5, 15, 23, 29, 35]
Year 10	3062	22.16 (8.31)	[5, 15, 23, 29, 35]

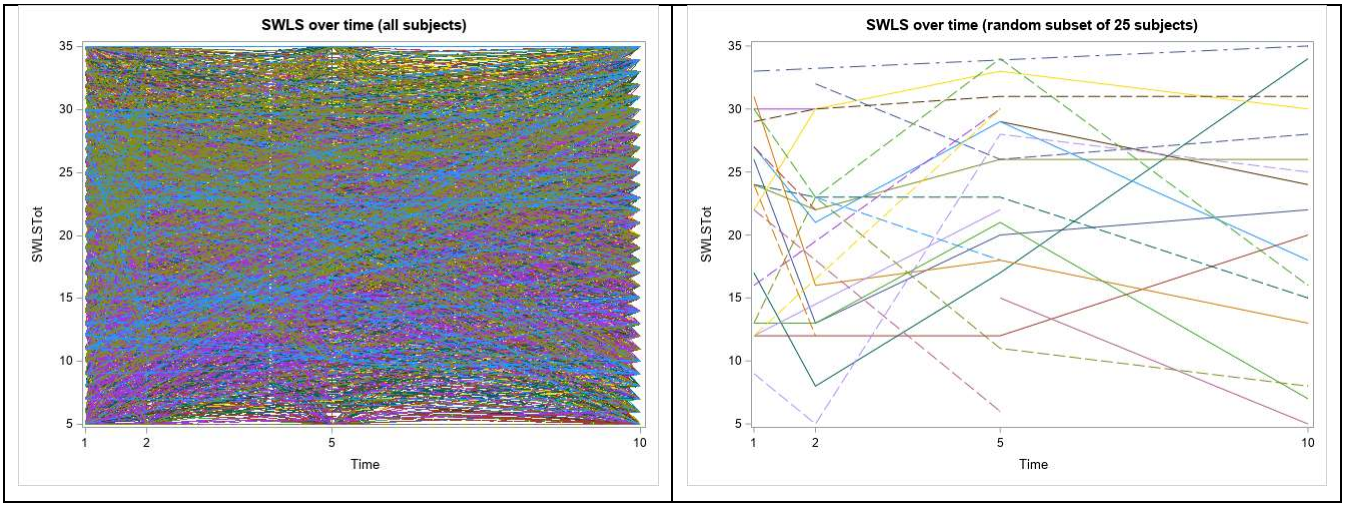
Pairwise correlations of the repeated SWLS outcomes are given in Table 2. There is variability in the correlations (0.50 – 0.64) over time suggesting a compound symmetric structure (requiring all these to be equal) may be too restrictive. The correlations follow the commonly seen pattern of diminishing in value as one goes further away from the diagonal. Repeated measures are less correlated the further away they are in time.

Table 2: Pairwise correlations in SWLS (N = 2550 – 3238); all significant at $p < 0.0001$

	Y1	Y2	Y5	Y10
Y1	1	0.64	0.54	0.50
Y2	0.64	1	0.62	0.54
Y5	0.54	0.62	1	0.62
Y10	0.50	0.54	0.62	1

The so called “spaghetti plots” of the data are presented in Figure 4. Each line represents the observed data from each individual, colored by center (left panel) or by subject (right panel). When dealing with a large number of subjects (even > 100), spaghetti plots of all subjects can be overwhelming and often look like an ink blotter as seen in the left panel of the Figure.

Figure 4: Spaghetti Plots for all subjects (left) colored by center and a random subset (right) colored by subject



This can be simplified by examining a random subset of the sample as shown in the right panel of the Figure. These plots can be useful in assessing the overall aspects of the data. For example, the spaghetti plots here suggest there is considerable heterogeneity between subjects in the initial SWLS at year 1, and the SWLS trajectory through year 10. Some subjects show increasing trajectories, some show decreasing trajectories, and others show initial decreases from year 1 to 2, followed by increased through years 5 and 10. It is helpful to keep an eye out for any meaningful trends observable by eye such as increasing/decreasing variance over time, pronounced linear or curvilinear trends over time, center or other clustering differences, and missing data patterns.

Example 1: Simple Linear Regression Model

We start this demonstration with the very basic simple linear regression model presented in equation (1). This is a simple linear regression model with only time as a regressor, where time is treated as a continuous variable taking on values of 1, 2, 5, and 10. This model does not include any random effects and assumes no correlation in the repeated measures over time. We present it here for completeness and reference. The estimated model parameters are summarized in Table 3.

Table 3: Simple Regression Model

Parameter	Estimate	SE	*Z*	*p*-value
β_0	21.11	0.11	183.89	< 0.0001
β_1	0.12	0.02	5.87	< 0.0001
σ^2	69.25	0.84		
-2LL	95463.9			
AIC	95465.9			

Note: -2LL = -2 Log Likelihood; AIC = Akaike Information Criteria = 2*k*-2LL, where *k* = number of model parameters; SE = Standard Error

If we assume no correlations, the intercept is estimated to be 21.11, and the linear slope is estimated to be 0.12. Thus, at time=0 SWLS scores are 21.11 and for each 1 year increase in time SWLS scores are estimated to increase by 0.12. As the range of time for our data is between 1 and 10 years, the intercept is not particularly meaningful (and certainly not interpretable as expected SWLS scores pre-injury or soon after injury). The estimate *is* meaningful as a part of the linear equation in (1) and can be interpreted as the underlying level of response in the population. It is also used to estimate means at each post injury year, along with the estimate of the slope. Specifically, $\hat{y}(t) = 21.11 + 0.12 \times t$, where *t* denotes year post injury. These estimated SWLS means at each year are summarized in Table 4 along with the observed SWLS means.

Table 4: Observed and Expected (Simple Linear Regression Model) Mean SWLS

	Year 1	Year 2	Year 5	Year 10
Observed	20.90	21.55	22.03	22.16
Estimated	21.23	21.36	21.72	22.30

A test of the intercept is also not meaningful, as it tests if the intercept (at time 0) is significantly different from 0; the *Z*-statistic and *p*-value are not interpreted here. A test of the slope parameter is meaningful, and it is our key parameter of interest. It test if the population change over time is different than 0. Here the slope of 0.12 is positive and significantly different than 0 ($p < 0.0001$). The residual error (assumed to be constant over time) is estimated to be 69.25 which is close to the variance (standard deviation squared) in SWLS scores observed across time (68.56-69.56).

Example 2: Random Intercepts MRM

Next, we demonstrate the random intercepts model corresponding to equation (2). We are extending the simple linear regression model in (1) by incorporating a random intercept term for each subject. Although our descriptive analysis suggested that a compound symmetric structure may be too simplistic, we will fit the random-intercept model here to demonstrate the extension from simple linear regression and before further extending it to account for more

complex modeling structures that may be more practical. The results from fitting this model using residual (restricted) maximum likelihood (REML) estimation are summarized in Table 5.

Table 5: Random Intercepts Model

Parameter	Estimate	SE	*Z*	*p*-value
β_0	20.96	0.12	168.06	< 0.0001
β_1	0.12	0.01	8.16	< 0.0001
$\sigma^2_{v_0}$	39.65	1.09		
σ^2	29.56	0.43		
-2LL	90839.9			
AIC	90843.9			

Note: -2LL = -2 Log Likelihood; AIC = Akaike Information Criteria = 2*k*-2LL, where *k* = number of model parameters; SE = Standard Error

Focusing first on the estimated regression parameters, this model estimates the population intercept as 20.96 and the slope as 0.12. Thus, at time=0 SWLS scores are estimated to be 20.96 and estimated to increase by 0.12 each year. The slope is statistically significant ($p < 0.0001$) and the total change in SWLS from 1 to 10 years is estimated to be about 1.4. The estimated intercept and slope are to calculate estimated values of mean SWLS at each timepoint. Specifically, $\hat{y}(t) = 20.96 + 0.12 \times t$. These are shown in Table 6 along with observed mean SWLS. Comparing the estimated to observe means in indicates the random intercept model provided a good fit of these observed means.

Table 6: Observed and Expected (Random Intercept Model) Mean SWLS

	Year 1	Year 2	Year 5	Year 10
Observed	20.90	21.55	22.03	22.16
Estimated	21.07	21.19	21.54	22.12

The model fit of the variances and covariances can also be examined. Here, the total estimated variance, which is assumed to be constant over time, is 39.65 + 29.56 = 69.20, or expressed as a standard deviation yields 8.32. Since the observed standard deviations in Table 1 do not change much over time (8.28-8.34), the estimate of a constant variance over time seems reasonable. Turning to the correlations of the repeated measures, the intraclass correlation here equals ICC=39.65/69.20 = 0.57, which indicates that 57% of the unexplained variance in SWLS scores (i.e., the part of SWLS not explained by the liner effect of time) is at the individual level (over half!). Thus, subjects display considerable heterogeneity in SWLS scores. Comparing this value of 0.57 to the observed correlation matrix in Table 2 suggests a good fit of the observed correlation structure.

Comparing the results from the random intercept model to the simple linear regression model we see that the regression parameter estimates are in close agreement, although their

standard errors are different. We note that what a simple linear regression lumped together into error variance (68.29 or 68.15), the random intercepts model separates into within subject (39.03) and between subject (29.13) variances.

In this model, we have allowed each individual to have their own deviation from the average trend in terms of the intercept term only. This assumed that the change over time was the same for all individuals; which is highly restrictive. In the next training module, we will extend the model to allow for each individual to have their own trend. We will also discuss how to recode time so that the intercept is more meaningful in the context of our example.

References

Berkhof, J. & Snijders, T.A.B. (2001). Variance component testing in multilevel models. *Journal of Educational and Behavioral Statistics*, 26, 133-152.

Bock, R.D. (1975). *Multivariate statistical methods in behavioral research*. New York: McGraw-Hill.

Bock, R.D. and Thissen, D., 1980, Statistical problems of fitting individual growth curves. In F.E. Johnston, A.F. Roche, and C. Susanne (Eds.), *Human Growth and Motivation: Methodologies and Factors*. New York: Plenum.

Bock, R.D. (1983a). The discrete Bayesian. In H. Wainer & S. Messick (Eds.), *Modern advances in psychiatric research* (pp. 103-115), Hillsdale, NJ: Lawrence Erlbaum.

Bock, R.D. (1983b). Within-subject experimentation in psychiatric research. In R.D Gibbons & M.W. Dysken (Eds.), *Statistical and methodological advances in psychiatric research* (pp. 59-90). New York: Spectrum.

Bock, R.D. (1989a). Measurement of human variations: A two stage model. In R.D. Bock (Ed.), *Multilevel analysis of educational data* (pp. 319-342). New York: Academic Press.

Bryk, A.S. & Raudenbush S.W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. In D. Berliner (Ed.), *Review of research in education* (Vol. 8, pp. 158-233). Washington, DC: American Educational Research Association.

de Leeuw, J. & Kreft, I. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics*, 11, 57-85.

Dempster, A.P., Rubin, D.B., & Tsutakawa, R.K. (1981). Estimation in covariance component models. *Journal of the American Statistical Society*, 76, 341-353.

Gibbons, R.D, Hedeker, D., Waternaux, C.M., & Davis, J.M. (1988(. Random Regression models: A comprehensive approach to the analysis of longitudinal psychiatric data. *Psychopharmacology Bulletin*, 244, 438-443.

Goldstein, H. (1981). *Measuring the stability of individual growth patterns*. Ann. Human Biology, 8: 549–557.

Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). New York: Halstead.

Hedeker (2004) Chapter 12: An Introduction to growth modeling, In D. Kaplan (Ed.) *Quantitative Methodology for Social Sciences*. Thousand Oaks CA: Sage.

Hedeker and Gibbons (2006) Chapters 4 and 5, Longitudinal Data Analysis, Wiley

Hui, S.L. & Berger, J.O. (1983). Empirical Bayes estimation of rates in longitudinal studies. *Journal of the American Statistical Association*, 78, 753-759.

Laird, N.M. & Ware, J.H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38, 963-974.

Longford, N.T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, 74, 817, 827.

Strenio, J.F., Weisberg, H.I., & Bryk, A.S. (1983). Empirical Bayes estimation of individual growth curve parameters and their relationship to covariates. *Biometrika*, 39, 71-86.

Verbeke, G. & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer.

Wald, A. (1943). Tests of hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54, 426-482.

Wolfinger, R.D. (1993), Covariance structure selection in general mixed models. *Communications in Statistics, Simulation and Computation*, 22, 1079-1106.