Modeling the Mean (I)

Biostatistics 653

Applied Statistics III: Longitudinal Data Analysis

General Guidelines

We introduce some general guidelines for selection of mean and covariance models for longitudinal data. Methods for model selection in longitudinal data are somewhat *ad hoc*, but we recommend the following:

- Select a preliminary mean structure. Because the covariance structure models all variability in the data that is not explained by fixed effects, we start by eliminating systematic trends. As recommended by Diggle (1994), we recommend using a rich mean structure, fitting a saturated model if possible, and a "largest reasonable" model otherwise.
- Select a preliminary covariance structure.
- Reduce the model if appropriate (concentrate on mean first, then check whether preliminary covariance structure is suitable for final mean structure). When evaluating the mean structure, be sure to use ML. ML or REML can be used to evaluate the covariance structure.
- Fit the final model using REML.

Modeling the Mean

When modeling the mean, we may consider some of the following parameterizations. We consider the case in which we have *H* groups and *n* repeated measures.

- *Profile Analysis*: This method may be used to describe treatment effects when there is no simpler parametric model and measurements are at a common set of times. It assumes we fit a separate mean for each group by time combination (with interaction). The hypothesis of no treatment effect corresponds to the test for no time by treatment interaction and has $(n-1)(H-1)$ degrees of freedom (df).
- *Parametric Curves*: These are useful when we expect changes to follow simple patterns over time and when measurements are not at a common set of times for all individuals. For example, we can treat time as a continuous variable and use a straight line model in time, in which case the test for no time by group interaction would have $(H-1)$ df.
- *Splines*: Splines are flexible functions that are useful when measurements are not at a common set of times and the change does not follow a simple pattern. These may require more specialized programming skills in Splus or R.

Modeling the Mean

- *Baseline or "Constant Effect" Model*: In some studies, an exposure or treatment might cause a location shift in the mean response that remains constant across measurement time.
- To fit a model that describes an effect that is constant over measurement occasions after baseline, we can create a new time variable: posttime=0 if baseline (t=0); posttime=1 if after baseline (t>0). Then we fit a model including group effects, posttime, and the group by posttime interaction.
- This model would test whether differences between group means, averaged over the n-1 post-baseline times of measurement, are significantly different from the corresponding group differences at baseline. This test is given by the posttime by group interaction test and has *H-1* df.
- We note that this model is nested in the saturated model and that it may be valid for randomized or observational studies.

Analyzing Response Profiles

- Consider a setting in which the design is balanced, with the timing of repeated measures the same for all individuals in the study (there may be some missing measurements).
- In the absence of information about the response trend over time, minimal restrictions can be made about the shape of the response trend by simply summarizing the estimated mean response at each occasion, stratified by levels of treatment or a group factor.
- At any given factor level, the sequence of means over time is called the mean response profile.
- While this approach may not be optimal for all analyses, it often represents a good starting point for selection of the preliminary mean structure of the model.

Analyzing Response Profiles

We focus on a two-group design, noting that generalizations to greater numbers of groups are straightforward. Given a sequence of *n* repeated measures on two groups of individuals, three main questions concerning the response profile may be of interest.

- Are the mean response profiles similar in the groups; that is, are they parallel? (group by time interaction)
- Assuming that the population mean response profiles are parallel, are they constant over time? (time effect)
- Assuming the population mean response profiles are parallel, are the profiles coincident? (group effect)

Randomized versus Observational Studies

- It is important to distinguish between longitudinal data coming from a randomized trial and data coming from an observational study.
- In a randomized study, when baseline values are obtained prior to study interventions (treatment), then we would expect (by design) for the group means to be equal at baseline. Thus if there is no interaction, there is no group effect.
- However, in an observational study, the groups will not necessarily have the same mean response at baseline (unless groups were selected by matching on baseline response).

Randomized Study: Lead Study

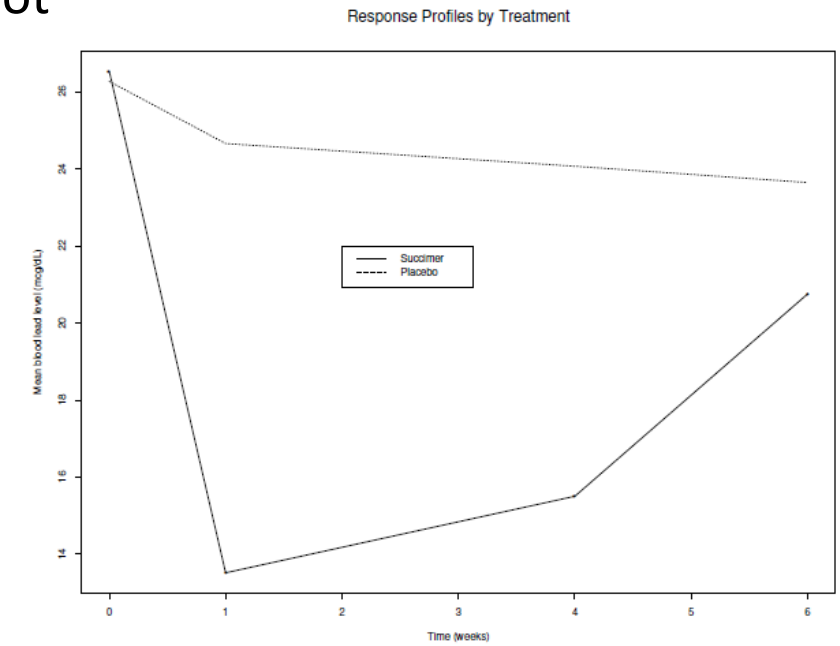
- Exposure to lead can cause cognitive impairment. Airborne lead levels have been dramatically reduced by the discontinuation of leaded gasoline; however, a small fraction of children are exposed to high levels of lead through deteriorating lead-based paint (present in many homes built before 1978). Lead poisoning in children can be treated by helping children excrete the ingested lead. A new chelating agent, succimer, enhances urinary excretion of lead and may be given orally (as opposed to older, injection-only treatments). A randomized clinical trial was conducted in children with confirmed high blood levels, who were randomized to receive either succimer or placebo and were followed longitudinally.

Lead Study

- In the lead study, the only question of interest is whether the response profiles are similar in the groups (question 1). Because the study is randomized, the profiles cannot be parallel if there is a treatment effect. The second question (whether the response profiles are constant over time) is of less importance, as it does not involve a direct comparison of the groups, and there is no interest whatsoever in the third question because the absence of a group by time interaction implies that there is no group effect in this randomized study.

Lead Study

- The mean proles at baseline and weeks 1, 4, and 6 of the study for subjects receiving treatment and placebo are presented in the following plot



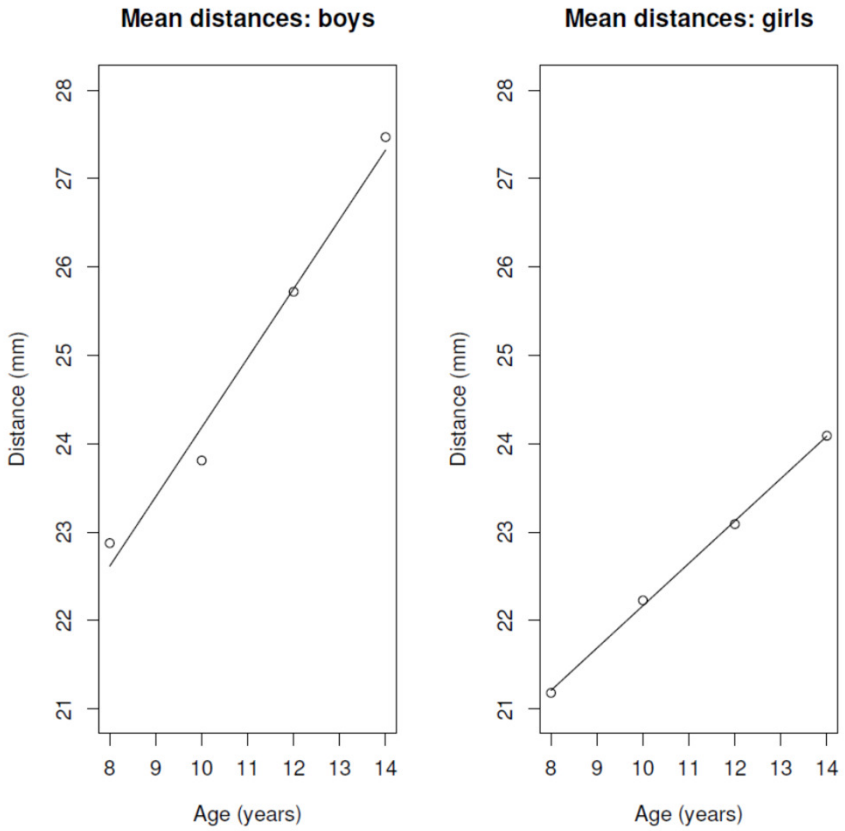
- While the observed baseline mean responses by group are not exactly the same, we would expect the population means to be the same (and certainly the estimated means are quite close to each other).

Observational Studies

- In an observational study, the first question (group by time interaction) is still of primary interest. However, the second and third questions may also be of interest. In particular, in this case, it is no longer true in general that the baseline responses must be similar, so that the absence of a group by time interaction does not imply the absence of a group effect.

Dental Study

- For example, consider the dental data. In this case, the boys had larger measurements than the girls at baseline.



Mean Model

- Consider the mean model $E(Y_i|X_i) = \mu_i = X_i\beta$. Suppose *n* is the number of repeated measures, and *N* is the number of study subjects. Also suppose that we have *H* groups of interest.
- For the dental study, N = 27, n = 4, and H = 2. Consider the cell means model

$$
\begin{aligned}
&E(Y_i|X_i) \\
&= \beta_1 I(boy, age\ 8) + \beta_2 I(boy, age\ 10) \\
&+ \beta_3 I(boy, age\ 12) + \beta_4\ I(boy, age\ 14) + \beta_5 I(girl, age\ 8) \\
&+ \beta_6 I(girl, age\ 10) + \beta_7 I(girl, age\ 12) + \beta_8\ I(girl, age\ 14)
\end{aligned}
$$

- This corresponds to the following design matrices

$$
\mathbf{X}_{i,boy} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \mathbf{X}_{i,girl} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}
$$

Mean Model

- For example, in this model, the mean response for boys at age 12 is given by β_3, and the mean response for girls at age 12 is given by β_7. Thus the gender difference at age 12 is given by $\beta_3 - \beta_7$.

- The hypothesis of no interaction (parallel profiles) is given by $H_{01}: (\beta_1 - \beta_5) = (\beta_2 - \beta_6) = (\beta_3 - \beta_7) = (\beta_4 - \beta_8)$, which is equivalent to $H_{01}: \boldsymbol{L\beta}$, where

$$\boldsymbol{L} = \begin{pmatrix} 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 \end{pmatrix}$$

so that

$$\boldsymbol{L\beta} = \begin{pmatrix} (\beta_1 - \beta_5) - (\beta_2 - \beta_6) \\ (\beta_1 - \beta_5) - (\beta_3 - \beta_7) \\ (\beta_1 - \beta_5) - (\beta_4 - \beta_8) \end{pmatrix}$$

Mean Model

- Instead of using cell mean coding, we could have used reference cell coding. In this case, we would have

$$\mathbf{X}_{i,boy} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{X}_{i,girl} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where girls at age 8 are the referent

- The coefficients are $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7)^T$, where α_0 is the intercept, $(\alpha_1, \alpha_2, \alpha_3)$ represent the main effects of age, α_4 is the gender indicator, and $(\alpha_5, \alpha_6, \alpha_7)$ represent the gender by age interaction terms.
- In this case a test of the hypothesis of no interaction is given by $H_{02}: \alpha_5 = \alpha_6 = \alpha_7 = 0$.

Covariance Structure

- In the analysis of response profiles, the covariance matrix is typically assumed to be unstructured. Multivariate Wald tests and likelihood ratio tests (using ML estimation) may be constructed for parameters of interest.
- Note that in the model selection tips given at the beginning of the lecture, we would begin by fitting the saturated model (preliminary mean structure) with unstructured covariance (preliminary covariance structure). We will later seek to refine the mean model and determine whether a more parsimonious model is adequate.

Group by Time Interaction

- The group by time interaction test in the dental study is given below in PROC MIXED. Note that we obtain this test in two ways: directly from the model output, and by specifying a specific contrast. In addition, we also illustrate a single degree of freedom test of whether the gender difference at age 14 is significant. In addition, note that we have set girls at age 8 to be the referent group in the first analysis, while in the second analysis, we use cell mean coding to make specification of the contrasts of interest more straightforward

Testing Contrasts

```
/* SORTING data in order to force age 8 to be referent */
proc sort data=proyuniv ; by descending time; run;
proc mixed data=proyuniv order=data;
class gender time newid;
model dist=time gender time*gender/ solution;
repeated/type=un subject=newid;
run;
/* now using cell mean coding for easy contrast estimation */
proc mixed data=proyuniv;
class gender time newid;
model dist=time*gender/noint solution;
repeated/type=un subject=newid;
contrast 'interaction test reproduced' time*gender 1 -1 0 0 -1 1 0 0,
        time*gender 1 0 -1 0 -1 0 1 0, time*gender 1 0 0 -1 -1 0 0 1/e chisq;
contrast 'test of difference in distance at age 14 by gender '
        time*gender 0 0 0 1 0 0 0 -1/e chisq;

run;
```

Contrast Estimation: Reference Cell

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Unstructured
Subject Effect	newid
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
GENDER	2	1 2
time	4	14 12 10 8
newid	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

Contrast Estimation: Reference Cell

Dimensions

Covariance Parameters	10
Columns in X	15
Columns in Z	0
Subjects	27
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	470.49084642	
1	1	414.03480100	0.00000000

Convergence criteria met.

Contrast Estimation: Reference Cell

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	newid	4.9857
UN(2,1)	newid	4.1307
UN(2,2)	newid	6.4557
UN(3,1)	newid	3.3172
UN(3,2)	newid	2.9272
UN(3,3)	newid	4.1848
UN(4,1)	newid	2.7102
UN(4,2)	newid	3.9102
UN(4,3)	newid	2.7168
UN(4,4)	newid	5.4155

Fit Statistics

-2 Res Log Likelihood	414.0
AIC (smaller is better)	434.0
AICC (smaller is better)	436.5
BIC (smaller is better)	447.0

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
9	56.46	<.0001

Contrast Estimation: Reference Cell

Solution for Fixed Effects

Effect	GENDER	time	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			21.1818	0.7017	25	30.19	<.0001
time		14	2.9091	0.6729	25	4.32	0.0002
time		12	1.9091	0.6068	25	3.15	0.0042
time		10	1.0455	0.6155	25	1.70	0.1018
time		8	0	.	.	.	.
GENDER	1		1.6932	0.9115	25	1.86	0.0750
GENDER	2		0	.	.	.	.
GENDER*time	1	14	1.6847	0.8741	25	1.93	0.0654
GENDER*time	1	12	0.9347	0.7883	25	1.19	0.2469
GENDER*time	1	10	-0.1080	0.7995	25	-0.14	0.8937
GENDER*time	1	8	0	.	.	.	.
GENDER*time	2	14	0	.	.	.	.
GENDER*time	2	12	0	.	.	.	.
GENDER*time	2	10	0	.	.	.	.
GENDER*time	2	8	0	.	.	.	.

Contrast Estimation: Reference Cell

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
time	3	25	34.45	<.0001
GENDER	1	25	9.29	0.0054
GENDER*time	3	25	2.93	0.0532

Contrast Estimation: Cell Mean

Model Information

Data Set	WORK.PROYUNIV
Dependent Variable	dist
Covariance Structure	Unstructured
Subject Effect	newid
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
GENDER	2	1 2
time	4	8 10 12 14
newid	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

Contrast Estimation: Cell Mean

Dimensions

Covariance Parameters	10
Columns in X	8
Columns in Z	0
Subjects	27
Max Obs Per Subject	4

Number of Observations

Number of Observations Read	108
Number of Observations Used	108
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	470.49084642	
1	1	414.03480100	0.00000000

Convergence criteria met.

Contrast Estimation: Cell Mean

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	newid	4.9857
UN(2,1)	newid	4.1307
UN(2,2)	newid	6.4557
UN(3,1)	newid	3.3172
UN(3,2)	newid	2.9272
UN(3,3)	newid	4.1848
UN(4,1)	newid	2.7102
UN(4,2)	newid	3.9102
UN(4,3)	newid	2.7168
UN(4,4)	newid	5.4155

Fit Statistics

-2 Res Log Likelihood	414.0
AIC (smaller is better)	434.0
AICC (smaller is better)	436.5
BIC (smaller is better)	447.0

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
9	56.46	<.0001

Contrast Estimation: Cell Mean

Solution for Fixed Effects

Effect	GENDER	time	Estimate	Standard Error	DF	t Value	Pr > |t|
GENDER*time	1	8	22.8750	0.5818	27	39.32	<.0001
GENDER*time	1	10	23.8125	0.5114	27	46.56	<.0001
GENDER*time	1	12	25.7187	0.6352	27	40.49	<.0001
GENDER*time	1	14	27.4687	0.5582	27	49.21	<.0001
GENDER*time	2	8	21.1818	0.7017	27	30.19	<.0001
GENDER*time	2	10	22.2273	0.6168	27	36.04	<.0001
GENDER*time	2	12	23.0909	0.7661	27	30.14	<.0001
GENDER*time	2	14	24.0909	0.6732	27	35.78	<.0001

Contrast Estimation: Cell Mean

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
GENDER*time	8	27	547.61	<.0001

Coefficients for interaction test replicated

Effect	GENDER	time	Row1	Row2	Row3
GENDER*time	1	8	1	1	1
GENDER*time	1	10	-1		
GENDER*time	1	12		-1	
GENDER*time	1	14			-1
GENDER*time	2	8	-1	-1	-1
GENDER*time	2	10	1		
GENDER*time	2	12		1	
GENDER*time	2	14			1

Contrast Estimation: Cell Mean

Coefficients for test of
diff in dist at age 14

Effect	GENDER	time	Row1
GENDER*time	1	8	
GENDER*time	1	10	
GENDER*time	1	12	
GENDER*time	1	14	1
GENDER*time	2	8	
GENDER*time	2	10	
GENDER*time	2	12	
GENDER*time	2	14	-1

Contrast Estimation: Cell Mean

Contrasts

Label	Num DF	Den DF	Chi-Square	F Value
interaction test replicated	3	27	8.79	2.93
test of diff in dist at age 14	1	27	14.92	14.92

Contrasts

Label	Pr > ChiSq	Pr > F
interaction test replicated	0.0322	0.0516
test of diff in dist at age 14	0.0001	0.0006

Other Tests of Group by Time Interaction

- In a typical randomized study, by design, the groups should have the same mean response at baseline. In such a setting, analysts may wish to specify a single degree of freedom contrast thought to best represent the direction in which the response pattern will differ the most.
- Suppose we wish to test the hypothesis that the difference between the average response at occasions 2 through *n* is equal to the difference in the baseline values in the two groups. Using cell mean coding in a study with two groups and four repeated measures, this corresponds to H_{03}: $\frac{(\beta_2+\beta_3+\beta_4)}{3} - \beta_1 = \frac{(\beta_6+\beta_7+\beta_8)}{3} - \beta_5$, where $L = \left(-1, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 1, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}\right)$.
- Why don't we always use this approach? This approach averages over the differences at all post-baseline occasions. This would not be a great strategy if we didn't expect to see much difference between the groups until the end of the study, for example.

Adjusting for Baseline Response

- In a randomized longitudinal study, the measurement at the rst occasion is usually a baseline response obtained prior to any study interventions. In that case, due to the randomization, we can assume that the treatment group means are equal at baseline. In this case, investigators may have questions about how to handle the baseline measure in the assessment of whether patterns of change in the mean response over time are the same between the two groups.

Adjusting for Baseline Response

We consider four methods of handling baseline response values.

- 1. Retain baseline response as part of the outcome vector, making no assumptions about group differences in mean response at baseline.
- 2. Retain baseline response as part of the outcome vector, assuming the group means are equal at baseline (typically appropriate in randomized study).
- 3. Subtract baseline response from all subsequent response measures, analyzing changes from baseline.
- 4. Use baseline as a covariate in analysis of subsequent responses.

Adjusting for Baseline Response

- Methods (1) and (2) both retain baseline as an outcome measure but make different assumptions about mean response at baseline. The second strategy constrains the baseline mean responses across groups to be equal (that is, eliminates the "succimer" main effect in reference cell coding), while the first one does not. (This is an important exception to the rule of not including interactions in a model without main effects.)
- In the context of a randomized study, both methods (1) and (2) are valid, though (2) is in general more powerful and is recommended as the standard analysis method. However, for observational studies, it is typically not the case that mean responses across all groups of interest are equal at baseline, so that method (2) is clearly inappropriate.

Adjusting for Baseline Response

- In the lead data, we consider the test of time by treatment interaction using methods (1) and (2).
- For method (1), we fit the model

$$
\begin{aligned}
&Y_{ij} \\
&= \beta_0 + \beta_1 I(week = 1)_{ij} + \beta_2 I(week = 4)_{ij} + \beta_3 I(week = 6)_{ij} \\
&+ \beta_4 I(succimer)_{ij} + \beta_5 I(succimer)_{ij} I(week = 1)_{ij} \\
&+ \beta_6 I(succimer)_{ij} I(week = 4)_{ij} \\
&+ \beta_7 I(succimer)_{ij} I(week = 6)_{ij} + \epsilon_{ij}
\end{aligned}
$$

- Where $\epsilon_i \sim MVN(0, \Sigma)$, with an unstructured covariance.
- To test the hypothesis of no interaction, we use

$$
L = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}
$$

Adjusting for Baseline Response

```
data new;
set lead;
time=0; lead=lead0; output;
time=1; lead=lead1; output;
time=4; lead=lead4; output;
time=6; lead=lead6; output;
drop lead0 lead1 lead4 lead6;
run;

data new;
set new;
succimer=0;
if trt="A" then succimer=1;
time0=0; time1=0; time4=0; time6=0;
if time=0 then time0=1;
if time=1 then time1=1;
if time=4 then time4=1;
if time=6 then time6=1;
run;
```

We obtain a Wald statistic of 109.99 on 3 degrees of freedom ($p < 0.0001$) for the test of time by treatment interaction.

Adjusting for Baseline Response

- Now, using the specication in (2), we t the model

$$
\begin{aligned}
&Y_{ij} \\
&= \beta_0 + \beta_1 I(week = 1)_{ij} + \beta_2 I(week = 4)_{ij} + \beta_3 I(week = 6)_{ij} \\
&+ \beta_4 I(succimer)_{ij} I(week = 1)_{ij} \\
&+ \beta_5 I(succimer)_{ij} I(week = 4)_{ij} \\
&+ \beta_6 I(succimer)_{ij} I(week = 6)_{ij} + \epsilon_{ij}
\end{aligned}
$$

- Where $\epsilon_i \sim MVN(0, \Sigma)$, with an unstructured covariance.

Adjusting for Baseline Response

```
proc mixed data=new noclprint method=ml;
class id;
model lead=time1 time4 time6 succimer*time1
        succimer*time4
        succimer*time6/solution;
repeated/type=un subject=id;
contrast 'interaction' succimer*time1 1,
        succimer*time4 1,
        succimer*time6 1/chisq e;
run;
```

We obtain a Wald statistic of 114.23 on 3 degrees of freedom ($p < 0.0001$) for the test of time by treatment interaction.

Adjusting for Baseline Response

- Methods (3) and (4) eliminate baseline response from the outcome vector.
- Method (3) involves construction of a new outcome variable, the change score, given by $D_i = (Y_{i2} - Y_{i1}, Y_{i3} - Y_{i1}, \cdots, Y_{in} - Y_{i1})^T$. Because the outcome is a change score, the interpretation of the typical tests in the analysis of response profiles changes. Suppose we are using reference cell coding. If we are interested in whether the patterns of change over time are the same in all H groups, we are no longer interested in just the group by time interaction test. (Think -- if the group by time interaction is not significant, but the step-down group test is significant, that means the change from baseline depends on group, which means the patterns of change over time are not the same for both groups).
- So to test that the patterns of change over time are the same in all H groups, we need to use a test of whether the main effect of group and the group by time interaction terms are equal to zero. It can be shown that this test is equivalent to the group by time interaction test in method (1).

Adjusting for Baseline Response

Though the methods (1) and (3) are equivalent for the purposes of evaluating the change in mean response over time, method (1) has two advantages:

- construction of hypothesis tests of interest are simpler using method (1), and
- when subjects are missing baseline response, all of their data are excluded from the analysis of change scores in method (3), while in method (1), all available data are used in analysis.

Adjusting for Baseline Response

- Method (4) treats baseline as a covariate and follow-up measures as the response of interest. This is appropriate for analyzing data from randomized trials. This method is also equivalent to treating baseline as a covariate and the change scores as the responses of interest (the only difference is in terms of the slope for the baseline response, which is one unit smaller in the change score analysis). Because of this correspondence, we can think of method (4) as an "adjusted change scores" approach, adjusting the approach in (3) for baseline response. So just as in method (3), method (4) requires a combined group and group by time interaction test in order to test the hypothesis that response profiles are the same over time in the groups of interest.

Adjusting for Baseline Response

- Methods (2) and (4) often give similar results. Method (2) is preferred for the same reasons method (1) is preferred over method (3). In addition, treating baseline as a covariate assumes the regression slope relating baseline to follow-up measures is the same at all follow-up occasions. This implies that $Cov(Y_{i1}, Y_{i2}) = Cov(Y_{i1}, Y_{i3}) = \cdots = Cov(Y_{i1}, Y_{ik})$. Because of this constraint, we could mis-specify the covariance model and make incorrect inferences about the pattern of change over time. Model (2) does not make this constraint, and thus may be more appropriate.

Adjusting for Baseline Response

- To summarize, for observational studies, the first method (retaining baseline in the outcome vector and making no assumptions about group means at baseline) is preferred. For randomized studies, the second method (retaining baseline in the outcome vector and assuming group means at baseline are equal) is typically more powerful.

Strengths of Analyzing Profiles

- Straightforward when design is balanced, timing of measures is common across individuals, and when all covariates are discrete
- Allows arbitrary patterns in mean response over time
- Allows arbitrary patterns in covariance of responses
- Thus has some robustness from risk of bias due to misspecification of mean and covariance models (makes minimal assumptions)

Weaknesses of Analyzing Profiles

- Not well-suited to handle measurements at different times across individuals
- Ignores time ordering of repeated measurements
- May have low power to detect group differences in specific trends over time (for example, testing linear trend in mean response over time)
- In saturated model, number of estimated mean and covariance parameters grows rapidly with number of measurement occasions (e.g., 12 parameters for 2 groups measured at 3 occasions; 75 parameters for 2 groups measured at 10 occasions!)