Advanced Biostatistics for Medical Researchers

Longitudinal Analysis

May, 2017

Approaches to Longitudinal Data Analysis

<u>References:</u>

- “Analysis of Longitudinal Data”. PJ Diggle, K-Y Liang, SL Zeger. 1994. Oxford University Press (Difficult)

- “ Statistical Methods for the Analysis of Repeated Measurement”. CS Davis. 2002. Springer-Verlag (Less Difficult)

- “Applied Longitudinal Analysis”. GM Fitzmaurice, NM Laird, JH Ware. 2004. John Wiley & Sons. (Best)

<u>Baseline CD4 Counts:</u> n=101, Mean=487 / Median=463, Range: [42,1093]

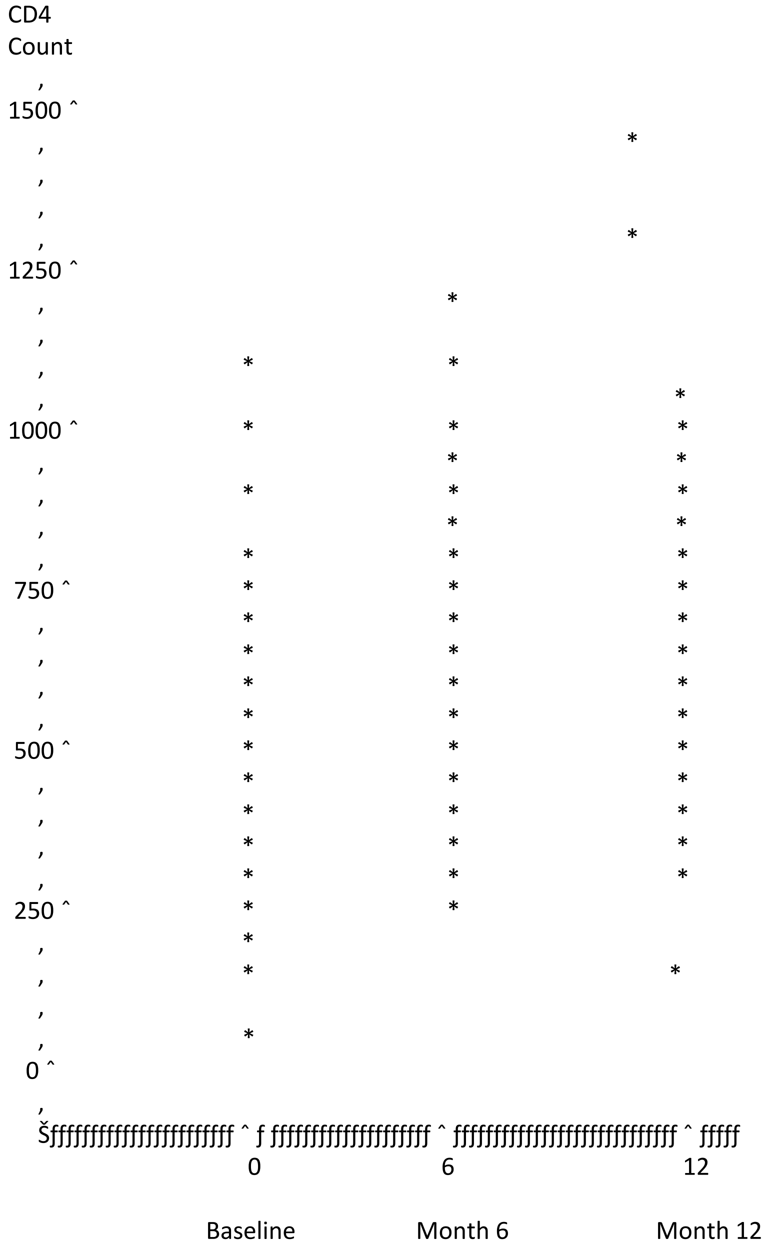
```
    ,                       *****
    ,                       *****
    ,                       *****
 25 ˆ                       *****     *****
    ,                       *****     *****
    ,                       *****     *****
    ,                       *****     *****
    ,                       *****     *****
 20 ˆ                       *****     *****
    ,                       *****     *****
    ,                       *****     *****
    ,             *****     *****     *****
    ,             *****     *****     *****
 15 ˆ             *****     *****     *****     *****
    ,             *****     *****     *****     *****
    ,             *****     *****     *****     *****
    ,             *****     *****     *****     *****
    ,             *****     *****     *****     *****
 10 ˆ             *****     *****     *****     *****
    ,             *****     *****     *****     *****
    ,             *****     *****     *****     *****
    ,             *****     *****     *****     *****
    ,             *****     *****     *****     *****     *****     *****
  5 ˆ             *****     *****     *****     *****     *****     *****
    ,             *****     *****     *****     *****     *****     *****
    ,   *****     *****     *****     *****     *****     *****     *****
    ,   *****     *****     *****     *****     *****     *****     *****
    ,   *****     *****     *****     *****     *****     *****     *****     *****
    Šƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
         75        225       375       525       675       825       975      1125
```

<u>6-Month CD4 Counts:</u> n=86, Mean=620 / Median=606, Range: [249,1213]

```
25 ˆ                         *****
   ,                         *****
   ,                         *****
   ,                         *****
   ,               *****     *****
20 ˆ               *****     *****
   ,               *****     *****
   ,               *****     *****
   ,               *****     *****
   ,               *****     *****
15 ˆ               *****     *****
   ,               *****     *****     *****
   ,               *****     *****     *****
   ,               *****     *****     *****
   ,     *****     *****     *****     *****     *****
10 ˆ     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****
 5 ˆ     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****     *****
   ,     *****     *****     *****     *****     *****     *****     *****
   Šƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
          300       450       600       750       900      1050      1200
```

12-Month CD4 Counts: n=80, Mean=695 / Median=683, Range: [168,1456]

```
    ,                 *****
    ,                 *****     *****
    ,                 *****     *****
    ,                 *****     *****
 20 ˆ                 *****     *****
    ,                 *****     *****
    ,                 *****     *****
    ,                 *****     *****
    ,       *****     *****     *****
 15 ˆ       *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
 10 ˆ       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
  5 ˆ       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    ,       *****     *****     *****     *****
    , *****     *****     *****     *****     *****     *****     *****
    Šƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
        200         400       600       800      1000      1200      1400
```


CD4
Count
1500
1250
1000
750
500
250
0
0
6
12
Baseline
Month 6
Month 12

Usual Approach for Independent Data: Linear Regression

$$CD4_i = \alpha + \beta\, Time_i + \varepsilon_i$$

where the subscript "i" goes across all of the measures of CD4 counts;

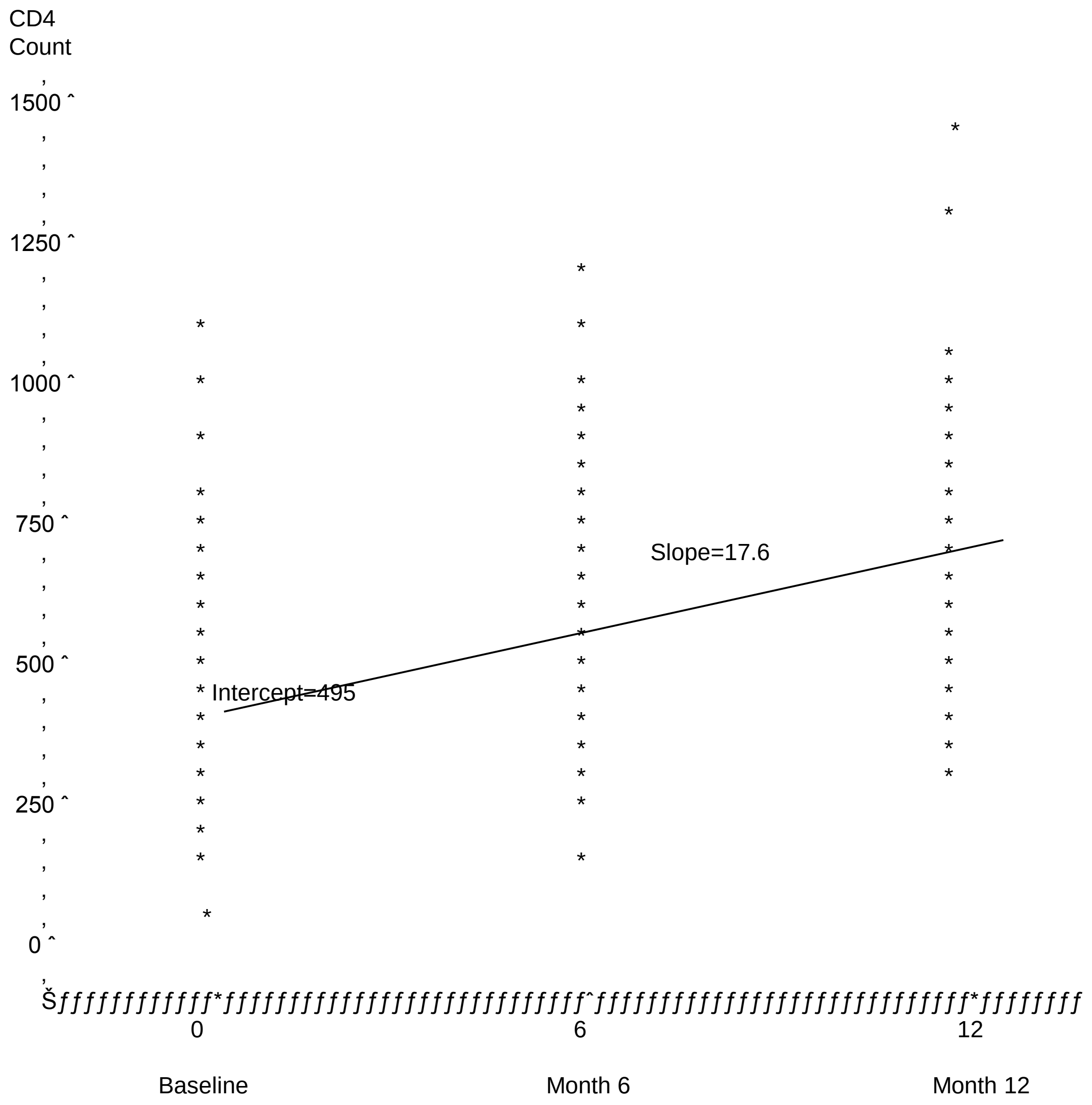
Assumption: The ε_i are independent of one another. → The CD4 counts are independent of one another.

Result:

CD4 = 495 + 17.6 x Time (in months) + Error

Where n = 267

Time effect = 17.6; se = 2.72; $p<.0001$


CD4
Count
1500
1250
1000
750
500
250
0
Slope=17.6
Intercept=495
0
6
12
Baseline
Month 6
Month 12

Problem: In this data set, there are multiple measurements of CD4 counts on each of 101 subjects.
(1 to 3 measurements per subject, taken at the start of treatment and then at 6 and 12 months)

→ We violate the independence assumption

More Appropriate Model:

$$CD4_{ij} = \alpha + \beta \, Time_{ij} + \varepsilon_{ij}$$

where “i” is an index for the 101 subjects (i.e., 1, 2, 3, … , 101)
and “j” is an index for serial measures on each subject
(i.e., j goes from 1 to 1, 2, or 3, depending on the subject)

Assumptions:

a. Measures of CD4 from different subjects are independent of one another.

b. Serial measures of CD4 from the same subject are correlated with each other.

→ **Repeated Measures Linear Regression Model**

(as opposed to a **Random Effects Model** or a **Random Coefficients Model**)

Comparisons Between Subjects
(with each subject having multiple correlated outcomes)

Basic Idea: If the CD4 counts are correlated within subject and, say, we want to compare CD4 counts between HIV-infected and HIV-uninfected subject (or, between treated and untreated; etc), then

a. 2 measurements on the same subject are not worth as much as 1 measurement on each of 2 subjects

or

b. 2 measurements on the same subject are not worth twice as much as one measurement

For example,

If Correlation=0, then	2 = 2x1
If Correlation=1, then	2 = 1
If 0<Correlation<1, then	2 = 1.?

i.e.,	If	Corr = .10,	then	2 = 1.8
	If	Corr = .25,	then	2 = 1.6
	If	Corr = .50,	then	2 = 1.3

Adjusting Sample Size Calculations For Correlated Data

Usual Situation: Common statistical tests and sample size calculations assume that the outcome measure acts independently from one measurement to the next. (If I told you the result of one measurement of the outcome, this would be totally uninformative about any other measurement of the outcome.)

This independence assumption is usually reasonable if the measurements are made only once on each subject, and the subjects have no relationship to each other.

This assumption is generally suspect in examples such as:

1. Each subject is measured at 4 time points.
2. The subjects are related (i.e., siblings or littermates)
3. Each subject contributes multiple measurements (i.e., digits, teeth, eyes, multiple segments of artery)

Situation of Current Interest:

- Assume that the experiment involves 25 subjects, but each subject is measured 4 times.
- This means that there are $4 \times 25 = 100$ outcome measurements available for analysis.
- However, the "effective sample size" is somewhere between N=25 (the number of subjects) and N=100 (the number of measurements).
- The true number depends on the magnitude of the correlation between the 4 measurements on each subject.

Definition 1: **(Intra-Class) Correlation Coefficient**

$$\rho = \sigma^2_{Between} \ / \ (\sigma^2_{Between} + \sigma^2_{Within})$$

(This ranges from 0 to 1, where we always presume positive correlation.)

Definition 2: **Design Effect (D)**

Assume that there are k correlated measures on each "subject"

$$D = 1 + (k-1)\rho$$

Result: Let N_{Random} represent the number of observations that would be required to carry out the study if all the observations were independent.

Then $$N_{Random} = N_{Correlated} / D$$

Prior Example: If we assume that for each of the 25 subjects, the correlation between the k=4 repeated measurements was ρ=0.5,

Then $$D = 1 + 3(.5) = 2.5$$

$$N_{Correlated} = 100$$

And $$N_{Random} = N_{Correlated} / D = 100/2.5 = 40$$

So, the "effective sample size" is 40

Example for Correlated Data

Sample Size Required, Per Group

Power=90%, Type I Error=5% (2-sided)

Standardized Effect Size = .50

	Number of Patients	Number of Measurements
Independent Data:	84	84
2 Measurements Per Subject		
ρ = .05	44	88
ρ = .10	46	92
ρ = .25	52	105
ρ = .50	63	126
3 Measurements Per Subject		
ρ = .05	31	92
ρ = .10	34	101
ρ = .25	42	126
ρ = .50	56	168
5 Measurements Per Subject		
ρ = .05	20	101
ρ = .10	24	118
ρ = .25	34	168
ρ = .50	50	252

Comparisons Within Subjects
(with each subject having multiple correlated outcomes)

Basic Idea: If the CD4 counts are correlated within subject and, say, we want to compare how CD4 counts change from baseline to follow-up within each subject (or, we do a cross-over study; etc), then:

a. 2 measurements on the same subject are worth **more** than 1 measurement on each of 2 subjects

b. 2 measurements on the same subject are worth **more** than twice as much as 1 measurement

For example,

If Correlation=0,	then 2 = 2x1
If Corr = .10,	then 2 = 2.2
If Corr = .25,	then 2 = 2.7
If Corr = .50,	then 2 = 4.0

As another example,

Assume we want to compare CD4 counts at diagnosis to CD4 counts 1-year after diagnosis, where CD4 counts have a standard deviation of 250 and change by 100 over the year.

a. If we measure 100 subjects at the time of diagnosis and a different 100 subjects at 1 year, we will have 80% power.
b. If we measure the same 100 subjects both at diagnosis and at 1 year, and assume a correlation of 0.30, we will have 92% power.

Concept: The Correlation Matrix

Assume that a subject has 5 measurements over time: Y_1, Y_2, Y_3, Y_4, and Y_5

And, let Corr(Y_1,Y_2) represent the correlation between Y_1 and Y_2, etc.

Then the **Correlation Matrix** represents all of the relationships between the five measurements:

	Y_1	Y_2	Y_3	Y_4	Y_5
Y_1	1.0	Corr(Y_1,Y_2)	Corr(Y_1,Y_3)	Corr(Y_1,Y_4)	Corr(Y_1,Y_5)
Y_2	Corr(Y_2,Y_1)	1.0	Corr(Y_2,Y_3)	Corr(Y_2,Y_4)	Corr(Y_2,Y_5)
Y_3	Corr(Y_3,Y_1)	Corr(Y_3,Y_2)	1.0	Corr(Y_3,Y_4)	Corr(Y_3,Y_5)
Y_4	Corr(Y_4,Y_1)	Corr(Y_4,Y_2)	Corr(Y_4,Y_3)	1.0	Corr(Y_4,Y_5)
Y_5	Corr(Y_5,Y_1)	Corr(Y_5,Y_1)	Corr(Y_5,Y_1)	Corr(Y_5,Y_1)	1.0

Note: We assume that every subject has the same correlation matrix

In Our Dataset: (Making no allowance for time trends)

	Baseline	6 Months	12 Months
Baseline	1.0	.53	.51
6 Months	.53	1.0	.69
12 Months	.51	.69	1.0

How Do We Estimate The Correlations?

We need to measure Corr(ε_{i1} , ε_{i2}), and Corr(ε_{i1} , ε_{i3}), and Corr(ε_{i2} , ε_{i3}), and … etc.

Which means we need to have the ε's, where

$$\varepsilon_{ij} = CD4_{ij} - \alpha - \beta \, Time_{ij}$$

Which means we need to have α and β.

Which means we have to run the regression.

But to run the regression we need to know how to weight the data, so we need to know the correlations.

→ This leads to a tedious, iterative fitting algorithm which is not always successful.

How Many Correlations Are We Talking About?
In our data, there are up to 3 measurements per subject, so there are only 3 unique correlation coefficients that need to be estimated.

However, if there were 15 serial measurements on a subject, there would be 105 unique correlation coefficients to be estimated.

Possible Simplifying Assumptions: We can help the process along by making some assumptions about the “form” of the correlation matrix. This will require the estimation of (many) fewer correlation coefficients, and lead to more stable answers concerning the regression parameters that are truly of interest.

1. Unstructured

	Y_1	Y_2	Y_3	Y_4	Y_5
Y_1	1.0	ρ_{12}	ρ_{13}	ρ_{14}	ρ_{15}
Y_2	ρ_{21}	1.0	ρ_{23}	ρ_{24}	ρ_{25}
Y_3	ρ_{31}	ρ_{32}	1.0	ρ_{34}	ρ_{35}
Y_4	ρ_{41}	ρ_{42}	ρ_{43}	1.0	ρ_{45}
Y_5	ρ_{51}	ρ_{52}	ρ_{53}	ρ_{54}	1.0

Toeplitz

	Y_1	Y_2	Y_3	Y_4	Y_5
Y_1	1.0	ρ_1	ρ_2	ρ_3	ρ_4
Y_2	ρ_1	1.0	ρ_1	ρ_2	ρ_3
Y_3	ρ_2	ρ_1	1.0	ρ_1	ρ_2
Y_4	ρ_3	ρ_2	ρ_1	1.0	ρ_1
Y_5	ρ_4	ρ_3	ρ_2	ρ_1	1.0

Autoregressive

	Y_1	Y_2	Y_3	Y_4	Y_5
Y_1	1.0	ρ	ρ^2	ρ^3	ρ^4
Y_2	ρ	1.0	ρ	ρ^2	ρ^3
Y_3	ρ^2	ρ	1.0	ρ	ρ^2
Y_4	ρ^3	ρ^2	ρ	1.0	ρ
Y_5	ρ^4	ρ^3	ρ^2	ρ	1.0

Exchangeable (Compound Symmetry)

	Y_1	Y_2	Y_3	Y_4	Y_5
Y_1	1.0	ρ	ρ	ρ	ρ
Y_2	ρ	1.0	ρ	ρ	ρ
Y_3	ρ	ρ	1.0	ρ	ρ
Y_4	ρ	ρ	ρ	1.0	ρ
Y_5	ρ	ρ	ρ	ρ	1.0

Data Analysis Results

Correlation Assumption	Monthly Change in CD4 Count (Slope)	Standard Error	P-value
1. Independence	17.6	2.72	p<.001
2. Exchangeable	17.6	1.84	p<.001
3. Autoregressive	18.0	2.18	p<.001
4. Toeplitz	17.8	1.98	p<.001
5. Unstructured	17.3	2.02	p<.001

Estimated Correlation Matrices

1. <u>Independence:</u>

	Baseline	6 Months	12 Months
Baseline	1.0	0.0	0.0
6 Months	0.0	1.0	0.0
12 Months	0.0	0.0	1.0

2. <u>Exchangeable:</u>

	Baseline	6 Months	12 Months
Baseline	1.0	0.56	0.56
6 Months	0.56	1.0	0.56
12 Months	0.56	0.56	1.0

3. <u>Autoregressive:</u>

	Baseline	6 Months	12 Months
Baseline	1.0	0.62	0.38
6 Months	0.62	1.0	0.62
12 Months	0.38	0.62	1.0

4. Toeplitz:

	Baseline	6 Months	12 Months
Baseline	1.0	0.61	0.49
6 Months	0.61	1.0	0.61
12 Months	0.49	0.61	1.0

5. Unstructured:

	Baseline	6 Months	12 Months
Baseline	1.0	0.53	0.49
6 Months	0.53	1.0	0.68
12 Months	0.49	0.68	1.0

An Alternative Way To Accommodate Correlation

Random Coefficients Model (Growth Curve)

$$CD4_{ij} = a_i + b_i\, Time_{ij} + \varepsilon_{ij}$$

Note: The slope and the intercept are now unique to each subject, rather than having one slope and one intercept which are representative of all subjects.

Assumptions: Across the subjects,

The slopes come from a Normal distribution, centered at β, with variance σ_β^2

These slopes are assumed to be independent from subject to subject. However, for a given subject, the same slope appears for each serial measurement and therefore there is correlation between serial measurements.

The intercepts come from a Normal distribution, centered at α, with variance σ_α^2

These intercepts are assumed to be independent from subject to subject. However, for a given subject, the same intercept appears for each serial measurement and therefore there is correlation between serial measurements.

The error terms (ε's) are assumed to be independent both within and between subjects

Data Analysis Results

Model Assumptions	Monthly Change in CD4 Count (Slope)	Standard Error	P-value
1. Random Intercepts	17.6	1.84	p<.001
2. Random Slopes	17.7	2.87	p<.001
3. Random Slopes and Intercepts	17.8	1.99	p<.001

Estimated Correlation Matrices

1. <u>Random Intercepts:</u>

	Baseline	6 Months	12 Months
Baseline	1.0	0.56	0.56
6 Months	0.56	1.0	0.56
12 Months	0.56	0.56	1.0

2. <u>Random Slopes:</u>

	Baseline	6 Months	12 Months
Baseline	1.0	0.0	0.0
6 Months	0.0	1.0	0.31
12 Months	0.0	0.31	1.0

3. <u>Random Slopes and Intercepts:</u>

	Baseline	6 Months	12 Months
Baseline	1.0	0.58	0.52
6 Months	0.58	1.0	0.64
12 Months	0.52	0.64	1.0

Obs	subject	cd4	time	Age	gender
1	JC-01*	1023	0	30.6438	M
2	JC-01*	1087	6	30.6438	M
3	JC-01*	651	12	30.6438	M
4	ND-02*	643	0	42.3507	M
5	ND-02*	561	6	42.3507	M
6	ND-02*	645	12	42.3507	M
7	SJ-03*	463	0	30.3315	M
8	SJ-03*	534	6	30.3315	M
9	SJ-03*	1019	12	30.3315	M
10	DK-04	652	0	34.9671	M
11	DK-04	777	6	34.9671	M
12	DK-04	682	12	34.9671	M
13	KM-05*	289	0	37.3753	M
14	KM-05*	608	6	37.3753	M
15	KM-05*	715	12	37.3753	M
16	KS-06*	551	0	36.1671	M
17	KS-06*	1085	6	36.1671	M
18	KS-06*	1037	12	36.1671	M
19	NF-07**	131	0	22.8767	F
20	NF-07**	548	6	22.8767	F
21	NF-07**	496	12	22.8767	F
22	RM-08*	263	0	39.1918	M
23	RM-08*	575	6	39.1918	M
24	RM-08*	631	12	39.1918	M
25	JF-09	365	0	35.6795	M
26	JF-09	459	6	35.6795	M
27	JF-09	594	12	35.6795	M
28	AR-10	919	0	30.6959	M
29	AR-10	757	6	30.6959	M
30	AR-10	836	12	30.6959	M
31	JB-11	311	0	30.8438	M
32	JB-11	605	6	30.8438	M
33	JB-11	506	12	30.8438	M
34	ACE-12	311	0	62.3342	M
35	ACE-12	473	6	62.3342	M
36	ACE-12	408	12	62.3342	M
37	JFM-13	667	0	32.4027	M
38	JFM-13	814	6	32.4027	M
39	JFM-13	769	12	32.4027	M
40	GV-14	981	0	44.6466	M
41	GV-14	973	6	44.6466	M
42	GV-14	1276	12	44.6466	M
43	BS-15*	413	0	40.6986	M
44	BS-15*	527	6	40.6986	M
45	BS-15*	555	12	40.6986	M
46	MB-16	1093	0	34.3315	M
47	MB-16	.	6	34.3315	M
48	MB-16	981	12	34.3315	M
49	SE-17	487	0	41.0329	M
50	SE-17	487	6	41.0329	M
51	SE-17	543	12	41.0329	M
52	RAS-18	303	0	38.4575	M
53	RAS-18	957	6	38.4575	M
54	RAS-18	406	12	38.4575	M
55	RC-19	443	0	55.7781	M

```
proc mixed;
 class subject;
 model  cd4  =  time / solution;
 repeated / type=ar(1) subject=subject rcorr;
```

```
proc mixed;
 class subject;
 model  cd4  =  time age / solution;
 repeated / type=toep subject=subject rcorr;
```

```
proc mixed;
 class subject gender;
 model cd4 = time age gender / solution;
 repeated / type=uns subject=subject rcorr;
```

The Mixed Procedure

Model Information

Data Set	WORK.LONGIDTUDINAL
Dependent Variable	cd4
Covariance Structure	Autoregressive
Subject Effect	subject
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
subject	103	ACE-12 AES-47 ALR-43 AP-22* AR-10 ASB-46* AV-35 BAB-25 BBB-125 BDD-56 BKC-32 BS-15* BSB-31 CCG-45 CP-52 CWM-102 DD-58 DIC-10 7 DJS-68 DK-04 DM-116 DMD-42 DRB-23 DT-90 DWS-85 EAS-80 ENS-41 EO-20 ESJ-27 FCHC 0135-78 FCHC 1026-64 FCHC 1029-67 FCHC 1031-77 FCHC 1033-7 3 FCHC 1045-99 FCHC 1049-111 FCHC 1051-112 FCHC1009-44 FCHC1010-40* FCHC1013-48 FD-96 FS-114 GAD-59 GI-88 GV-14 IC-51 JB-11 JC-01* JDB-74 JF-09 JFM-13 JGN-100 JH-33 JLS-70 JMV-63 JRP-54 JXV-84 KH -105 KJS-66* KM-05* KS-06* LFD-39 MB-16 MCW-29 MDL-81 ME-75 MJH-72 MJK-95 ML-37 MLD-86 MLL-61* MRB-62 MTF-76 ND-02* NEC-119 NF-07** NJB-28 NTL-106 PCF-26 PS-87 RAB-117 RAS-18 RC-19 RDS-103 REM-101 RJ-94 RJS-30 RLO-113 RM-08* RSN-91 SE-17 SG-34 SJ-03* SO-65 SRC-21 TD-36 TNT-55 WB-24 WF-104 WJO-82 WPC-69 WRC-60 WSB-118

Dimensions

Covariance Parameters	2
Columns in X	2
Columns in Z	0
Subjects	103
Max Obs Per Subject	3

Number of Observations

Number of Observations Read	309
Number of Observations Used	267
Number of Observations Not Used	42

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	3620.75413134	
1	2	3547.30774003	0.00000666
2	1	3547.29745747	0.00000000

Convergence criteria met.

Estimated R Correlation Matrix
for subject ACE-12

Row	Col1	Col2	Col3
1	1.0000	0.6186	0.3826
2	0.6186	1.0000	0.6186
3	0.3826	0.6186	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
AR(1)	subject	0.6186
Residual		48563

Fit Statistics

-2 Res Log Likelihood	3547.3
AIC (smaller is better)	3551.3
AICC (smaller is better)	3551.3
BIC (smaller is better)	3556.6

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
1	73.46	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	491.27	21.8234	100	22.51	<.0001
time	18.0256	2.1757	165	8.29	<.0001

The Mixed Procedure

Model Information

Data Set	WORK.LONGIDTUDINAL
Dependent Variable	cd4
Covariance Structure	Toeplitz
Subject Effect	subject
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
subject	103	ACE-12 AES-47 ALR-43 AP-22* AR-10 ASB-46* AV-35 BAB-25 BBB-125 BDD-56 BKC-32 BS-15* BSB-31 CCG-45 CP-52 CWM-102 DD-58 DIC-107 DJS-68 DK-04 DM-116 DMD-42 DRB-23 DT-90 DWS-85 EAS-80 ENS-41 EO-20 ESJ-27 FCHC 0135-78 FCHC 1026-64 FCHC 1029-67 FCHC 1031-77 FCHC 1033-73 FCHC 1045-99 FCHC 1049-111 FCHC 1051-112 FCHC1009-44 FCHC1010-40* FCHC1013-48 FD-96 FS-114 GAD-59 GI-88 GV-14 IC-51 JB-11 JC-01* JDB-74 JF-09 JFM-13 JGN-100 JH-33 JLS-70 JMV-63 JRP-54 JXV-84 KH -105 KJS-66* KM-05* KS-06* LFD-39 MB-16 MCW-29 MDL-81 ME-75 MJH-72 MJK-95 ML-37 MLD-86 MLL-61* MRB-62 MTF-76 ND-02* NEC-119 NF-07** NJB-28 NTL-106 PCF-26 PS-87 RAB-117 RAS-18 RC-19 RDS-103 REM-101 RJ-94 RJS-30 RLO-113 RM-08* RSN-91 SE-17 SG-34 SJ-03* SO-65 SRC-21 TD-36 TNT-55 WB-24 WF-104 WJO-82 WPC-69 WRC-60 WSB-118

Dimensions

Covariance Parameters	3
Columns in X	3
Columns in Z	0
Subjects	103
Max Obs Per Subject	3

Number of Observations

Number of Observations Read	309
Number of Observations Used	267
Number of Observations Not Used	42

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	3617.79552161	
1	2	3541.26198379	0.00014941
2	1	3541.01796083	0.00000206
3	1	3541.01478757	0.00000000

Convergence criteria met.

Estimated R Correlation Matrix
for subject ACE-12

Row	Col1	Col2	Col3
1	1.0000	0.6122	0.4935
2	0.6122	1.0000	0.6122
3	0.4935	0.6122	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
TOEP(2)	subject	29651
TOEP(3)	subject	23902
Residual		48433

Fit Statistics

-2 Res Log Likelihood	3541.0
AIC (smaller is better)	3547.0
AICC (smaller is better)	3547.1
BIC (smaller is better)	3554.9

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
2	76.78	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	474.41	83.2257	99	5.70	<.0001
time	17.8324	1.9851	165	8.98	<.0001
Age	0.5355	2.1644	99	0.25	0.8051

The Mixed Procedure

Model Information

Data Set	WORK.LONGIDTUDINAL
Dependent Variable	cd4
Covariance Structure	Unstructured
Subject Effect	subject
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
subject	103	ACE-12 AES-47 ALR-43 AP-22* AR-10 ASB-46* AV-35 BAB-25 BBB-125 BDD-56 BKC-32 BS-15* BSB-31 CCG-45 CP-52 CWM-102 DD-58 DIC-107 DJS-68 DK-04 DM-116 DMD-42 DRB-23 DT-90 DWS-85 EAS-80 ENS-41 EO-20 ESJ-27 FCHC 0135-78 FCHC 1026-64 FCHC 1029-67 FCHC 1031-77 FCHC 1033-73 FCHC 1045-99 FCHC 1049-111 FCHC 1051-112 FCHC1009-44 FCHC1010-40* FCHC1013-48 FD-96 FS-114 GAD-59 GI-88 GV-14 IC-51 JB-11 JC-01* JDB-74 JF-09 JFM-13 JGN-100 JH-33 JLS-70 JMV-63 JRP-54 JXV-84 KH -105 KJS-66* KM-05* KS-06* LFD-39 MB-16 MCW-29 MDL-81 ME-75 MJH-72 MJK-95 ML-37 MLD-86 MLL-61* MRB-62 MTF-76 ND-02* NEC-119 NF-07** NJB-28 NTL-106 PCF-26 PS-87 RAB-117 RAS-18 RC-19 RDS-103 REM-101 RJ-94 RJS-30 RLO-113 RM-08* RSN-91 SE-17 SG-34 SJ-03* SO-65 SRC-21 TD-36 TNT-55 WB-24 WF-104 WJO-82 WPC-69 WRC-60 WSB-118
gender	2	F M

Dimensions

Covariance Parameters	6
Columns in X	5
Columns in Z	0
Subjects	103
Max Obs Per Subject	3

Number of Observations

Number of Observations Read	309
Number of Observations Used	267
Number of Observations Not Used	42

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	3607.83497574	
1	2	3525.89724948	0.00000538
2	1	3525.88893623	0.00000000

Convergence criteria met.

Estimated R Correlation Matrix
for subject ACE-12

Row	Col1	Col2	Col3
1	1.0000	0.5352	0.5013
2	0.5352	1.0000	0.6831
3	0.5013	0.6831	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	subject	48169
UN(2,1)	subject	24333
UN(2,2)	subject	42911
UN(3,1)	subject	25606
UN(3,2)	subject	32930
UN(3,3)	subject	54158

Fit Statistics

-2 Res Log Likelihood	3525.9
AIC (smaller is better)	3537.9
AICC (smaller is better)	3538.2
BIC (smaller is better)	3553.7

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
5	81.95	<.0001

Solution for Fixed Effects

Effect	gender	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		482.96	84.5738	98	5.71	<.0001
time		17.2929	2.0208	98	8.56	<.0001
Age		0.4622	2.1803	98	0.21	0.8326
gender	F	18.9139	79.1864	98	0.24	0.8117
gender	M	0	.	.	.	.