Multilevel Models Workshop

Cindy L. Christiansen

Susan Loveland

CHQOER at Bedford VA and Boston University School of Public Health

Bedford Center for Health Quality, Outcomes, & Economics Research

WHEN DO YOU NEED A MULTILEVEL MODEL ?

- Longitudinal or Repeated Measures
    - Multiple measurements per subject or unit

- Nested or Clustered Data
    - patients within hospitals
    - hospitals within VISNs

Multilevel Models by any other name…

- Hierarchical Models
- Longitudinal Models
- Growth Curve Models
- Bayesian Models
- Random Effects Models
- Latent Variable Models

DATA and PARAMETERS

- DATA = anything that you need in your data file to use the software package and selected model
  – Outcome or measure (dependent variable {Level 1})
  – Independent variables or control variables (age, female, disease status, …)
  – Cluster variable (patient, hospital, …{defines Level 2})
  – Indicator of time of measure (for repeated measures)

DATA and PARAMETERS

- PARAMETERS = true population values; We need ESTIMATES of parameters. Parameters have DATA connected to them
  - Beta coefficients, regression coefficients, fixed effects (Independent variables: age, female, disease status, …)
  - Random effect parameters, variance parameters (Cluster variable: patient, hospital, …)
  - Slope or curve parameter (Indicator of time of measure)

FIXED AND RANDOM EFFECTS?

- **Why I don't use the term "fixed and random effects", Andrew Gelman, 1/25/05** http://www.stat.columbia.edu/~cook/movabletype/archives/2005/01/why_i_dont_use.html
  - "No agreed-upon meaning"
  - "I'm almost sure SAS uses definition (5)." "(5) Fixed effects are estimated using … maximum likelihood and random effects are estimated with shrinkage …"
- CLC disagrees! I think SAS uses definition "(1) Fixed effects are constant across individuals, and random effects vary." (vary across hospitals in our case)

And to make things more confusing…What is a Level?

- Some multilevel models have DATA (and PARAMETERS) at more than 1 level
  - patient mental health score & teaching/non-teaching status of hospitals
- Some multilevel models have PARAMETERS at Levels 2 or higher but no DATA at these levels except the cluster variable
  - Intercept and slope parameters at the patient level & random effect parameters for hospital intercepts

Examples of Multilevel Models shown today…

- Multilevel model with
  - Patients within hospitals
  - Outcome Measure: mental health score that is typically around 100 with a SD of 10
  - Independent Variable: age
  - Level 1 = patient; Level 2 = hospital (SAS calls this level "subjects")

References for help fitting other types of models provided in the handout.

DATA and PARAMETERS for our Examples

- Data:
  - Mental health score (MHS) for 2565 patients at 100 hospitals (dependent variable)
  - Age of patient (independent variable)
  - Hospital indicator (to define clusters)
- Parameters – What are we interested in?
  - Effect of hospitals on MHS
    - Random effect of hospitals (does the intercept vary by hospitals?)
    - Variance of random effect (how much do the intercepts vary?)
  - Effect of age on the MHS
    - Fixed effect (Beta coefficient of age; slope)
    - Random effect (does the effect of age (slope) vary across hospitals?)

SO MANY PROCS – SO LITTLE TIME

SAS

- Proc GLM
- Proc Logistic
- Proc Mixed
- Proc GLIMMIX
- Proc NLMIXED
- Proc Genmod

DATA

10 of 2,565 obs

Obs	MHS	age	centered_age (=age-65)	sta *
1	107.45767989	62	-3	1
2	113.39441701	66	1	1
3	99.903304035	64	-1	1
4	105.34759809	64	-1	1
5	112.27412702	69	4	1
6	97.298690021	63	-2	1
7	110.35388017	66	1	1
8	111.54640742	64	-1	1
9	101.73315717	66	1	1
10	100.77299724	65	0	1

* sta ranges from 1 to 100

PROC GLM VS MIXED

Fixed Effects

```
PROC GLM DATA=in.data;
MODEL MHS=AGE/solution;
RUN;

PROC MIXED DATA=in.data;
MODEL MHS=AGE/solution;
RUN;
```

GLM OUTPUT
(remember FIXED) (Page 2)

the GLM Procedure

Dependent Variable: MHS

Source	DF	Type III SS	Mean Square	F Value	Pr > F
age	1	884.9865548	884.9865548	7.15	0.0076

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	81.26893187	6.96671265	11.67	<.0001
age	0.28656649	0.10719438	2.67	0.0076

MIXED OUTPUT
(remember FIXED) (Page 1)

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	81.2689	6.9667	2563	11.67	<.0001
age	0.2866	0.1072	2563	2.67	0.0076

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
age	1	2563	7.15	0.0076

MIXED + GLM OUTPUT
(remember FIXED) (Pages 1 & 2)

Dependent Variable: MHS - GLM

Parameter	Estimate	StdError	t Value	Pr > |t|
Intercept	81.268931	6.96671265	11.67	<.0001
age	0.286566	0.10719438	2.67	0.0076

****

****

****

Solution for Fixed Effects (MIXED)

Effect	Estimate	StdError	DF	t Value	Pr > |t|
Intercept	81.2689	6.9667	2563	11.67	<.0001
age	0.2866	0.1072	2563	2.67	0.0076

We haven’t done anything “MULTILEVEL” yet!

Unconditional and Conditional Models – MORE TERMS!!!

(Page 3; Unconditional Model)

```
PROC mixed DATA=in.data covtest;
MODEL MHS = /solution;
```

(Page 1; Conditional Model)

```
PROC MIXED DATA=in.data covtest;
MODEL MHS=AGE/solution;
```

We still aren't doing anything "Multilevel"!

Unconditional & Conditional Models

```
(Page 3; Unconditional Model)
PROC mixed DATA=in.data covtest;
MODEL MHS = /solution;

Solution for Fixed Effects
Effect      Estimate   Std Error   DF    tValue   Pr>|t|
Intercept   99.8840    0.2200     2564    454.05  <.0001

(Page 1; Conditional Model)
PROC MIXED DATA=in.data covtest;
MODEL MHS=AGE/solution;

Solution for Fixed Effects
Effect    Estimate    std Error   DF    tValue   Pr>|t|
Intercept 81.2689     6.9667     2563  11.67   <.0001
age         0.2866    0.1072     2563   2.67   0.0076

We still aren't doing anything "Multilevel"!
```

Effect of Age on the Residual Variance

(Page 3; Unconditional Model)

```
MODEL MHS = /solution

Covariance Parameter Estimates
Cov Parm   Estimate   Std Error   Z Value    Pr Z
Residual    124.13     3.4668      35.81   <.0001
```

(Page 1; Conditional Model)

```
MODEL MHS=AGE/solution

Covariance Parameter Estimates
Cov Parm   Estimate   Std Error   Z Value     Pr Z
Residual     123.83   3.4592      35.80     <.0001
********
```

So, AGE accounts for (124.13-123.83)/124.13 of the variance in MHS's (almost nothing!)

What is centering?

How do we interpret results with or without centering?

(Page 1)

```
MODEL MHS=AGE/solution
```

Solution for Fixed Effects

Effect	Estimate	std Error	DF	tValue	Pr>|t|
Intercept	81.2689	6.9667	2563	11.67	<.0001
age	0.2866	0.1072	2563	2.67	0.0076

(Page 6)

```
MODEL MHS = centered_age/solution
```

Solution for Fixed Effects

Effect	Estimate	Std Error	DF	tValue	Pr>|t|
Intercept	99.8958	0.2198	2563	454.56	<.0001
centered_age	0.2866	0.1072	2563	2.67	0.0076

3 Things to Remember about Fitting Data to Multilevel Models

1. Look at your data – make eye contact
   - Check the # of clusters
   - Check the range of sample size across clusters
   - Be creative - Graph something!
2. Center independent variables; let the reference group represent a typical person
   - Important for interpretation
   - Important for convergence (be nice to your software!)
3. Model fitting is ITERATIVE – particularly for Multilevel Models
   - you will have to fit more than 1 model before you are finished!

Let's Do Multilevel Models

STA (Hospital) = LEVEL 2, Random Intercepts

Patient=LEVEL 1

AGE=FIXED, LEVEL 1 Parameter

(Page 4)

```
CLASS sta; MODEL
  MHS=age/solution;
RANDOM int / subject=sta;
```

What does it mean to have random intercept for hospitals?

(Page 4)

CLASS sta; MODEL MHS=age/solution; RANDOM int / subject=sta;

```
The Mixed Procedure
          Model Information
Data Set                        IN. DATA
Dependent Variable              MHS
Covariance Structure            Unstructured
Subject Effect                  sta
Estimation Method               REML
Residual Variance Method        Profile
Fixed Effects SE Method         Model-Based
Degrees of Freedom Method       Between-Within
```

(Page 4)

```
CLASS sta; MODEL MHS=age/solution; RANDOM int /sub=sta;
```

Dimensions

Covariance Parameters	2
Columns in X	2
Columns in Z Per Subject	1
Subjects	100
Max Obs Per Subject	50

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	19641.50320947	
1	2	19307.39611895	0.00000003
2	1	19307.39591948	0.00000000

Convergence criteria met.

(Page 4)

```
CLASS sta; MODEL MHS=age/solution; RANDOM int /sub=sta;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Std Error	Z Value	Pr Z
UN(1,1)	sta	22.53	4.0028	5.63	<.0001
Residual		101.62	2.8928	35.13	<.0001

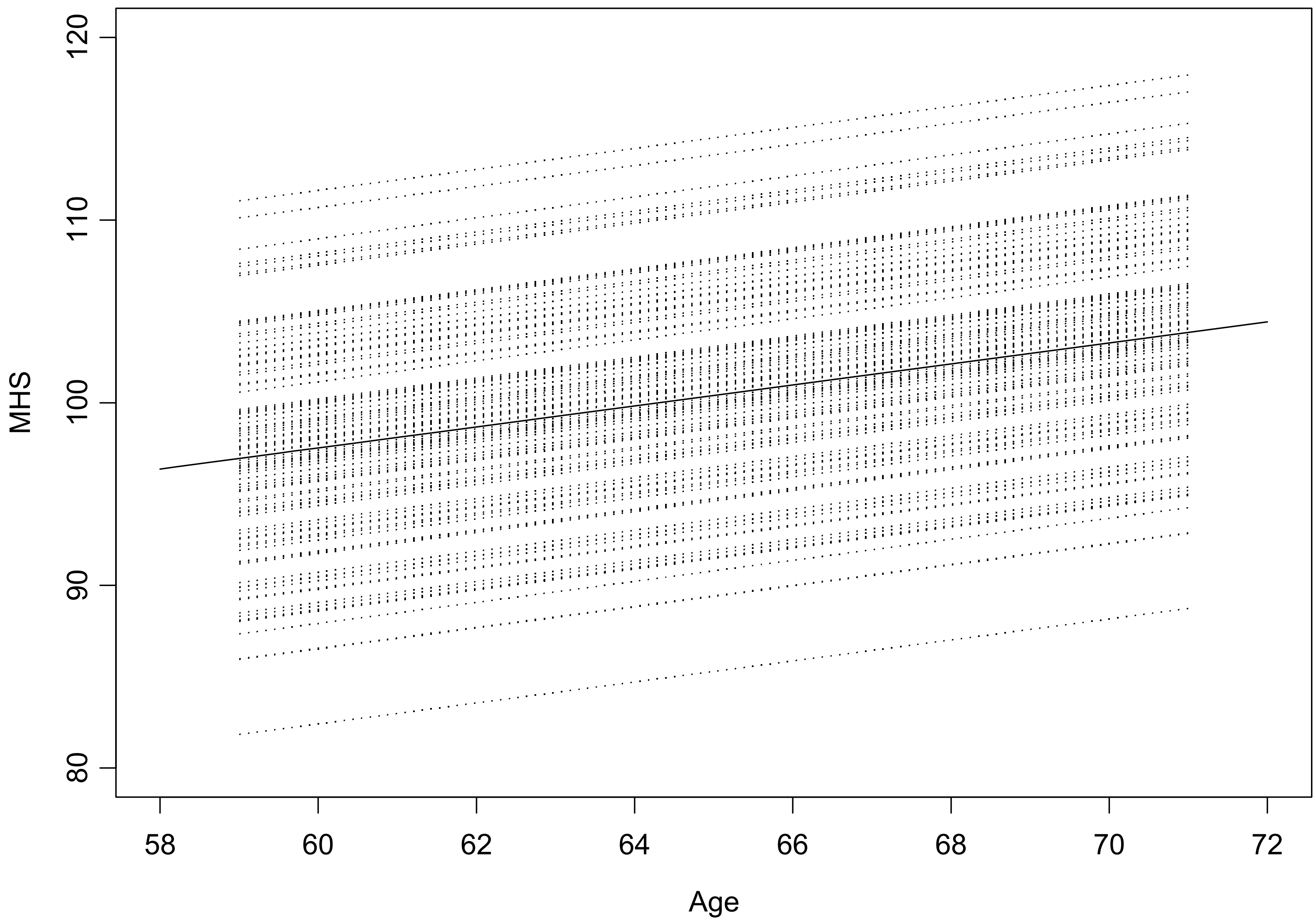
Solution for Fixed Effects

Effect	Estimate	std Error	DF	tValue	Pr>|t|
Intercept	78.7023	6.4190	99	12.26	<.0001
age	0.3222	0.09846	2464	3.27	0.0011

Group Discussion: Pages 3, 6 thru 8

Example 1		Unconditional Page 3	Conditional Page 6	Random Intercepts Page 7	Random Intercepts & Slopes Page 8
Fixed Effects (Level 1)	Intercept (age centered at 65)	99.9	99.9	99.6	99.6
	Centered Age		.29 (.11)	.32 (.10)	.32 (.10)
Level 1	Residual	124.1	123.8	101.6	101.6
Random Effects (Level 2)	(1,1) Intercepts			22.5	22.5
	(1,2)				.15
	(2,2) Slopes				0
	AIC	19648	19644	19311	19313
	-2LL	19646	19642	19307	19307

Predictions of the effect of Age on MHS by hospitals (dotted lines) and overall (solid line) (model on Page 7)

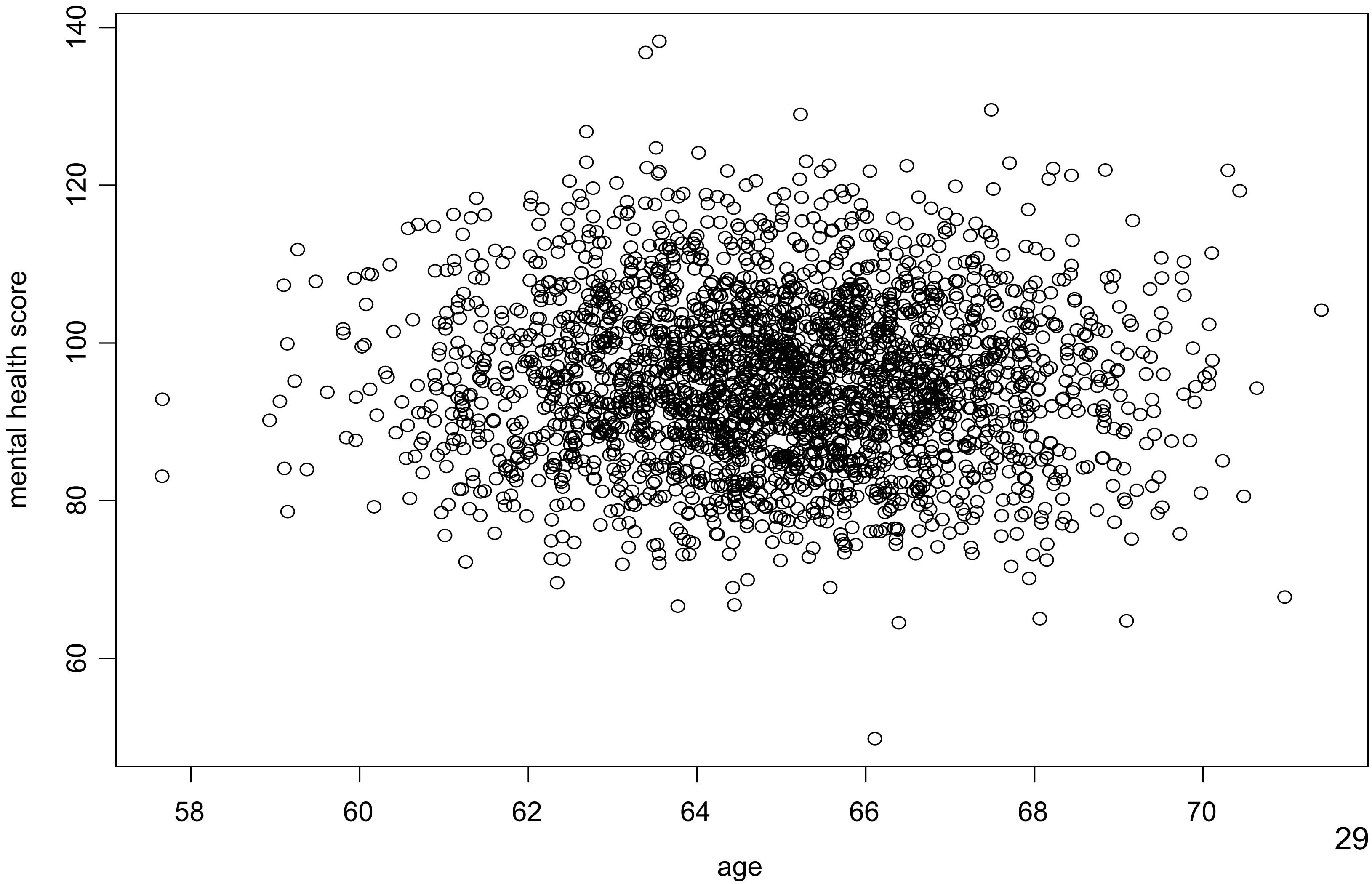


Data for Example 2
2565 observations at 100 hospitals

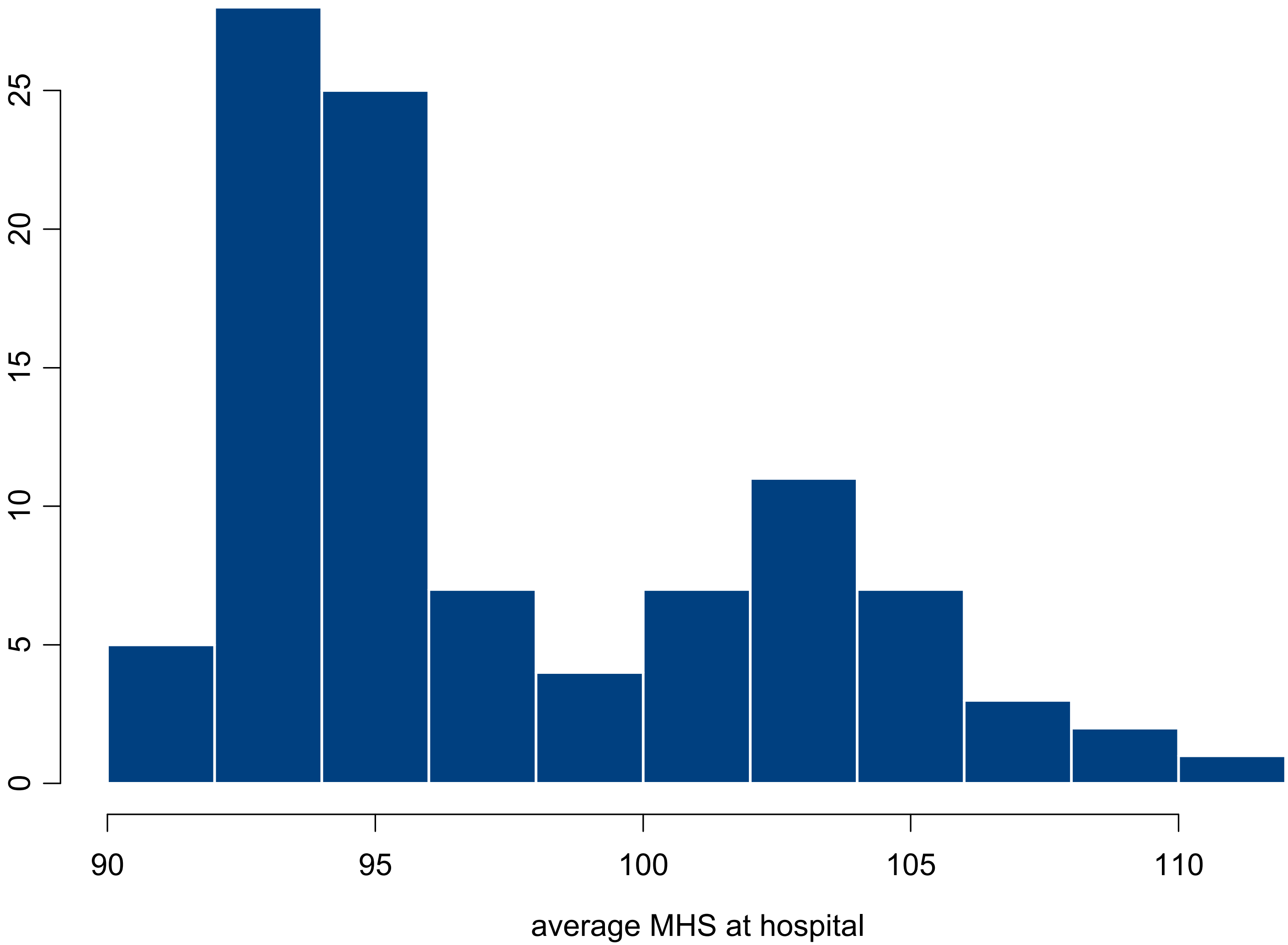
	# obs per hospital*	Mental Health Score	Age*
Minimum	1	49.8	58
1st Quartile	13	88.3	64
Median	26	95.1	65
Mean	25.7	95.3	65
3rd Quartile	38	102.2	66
Maximum	50	138.3	71

* Same as in Example #1

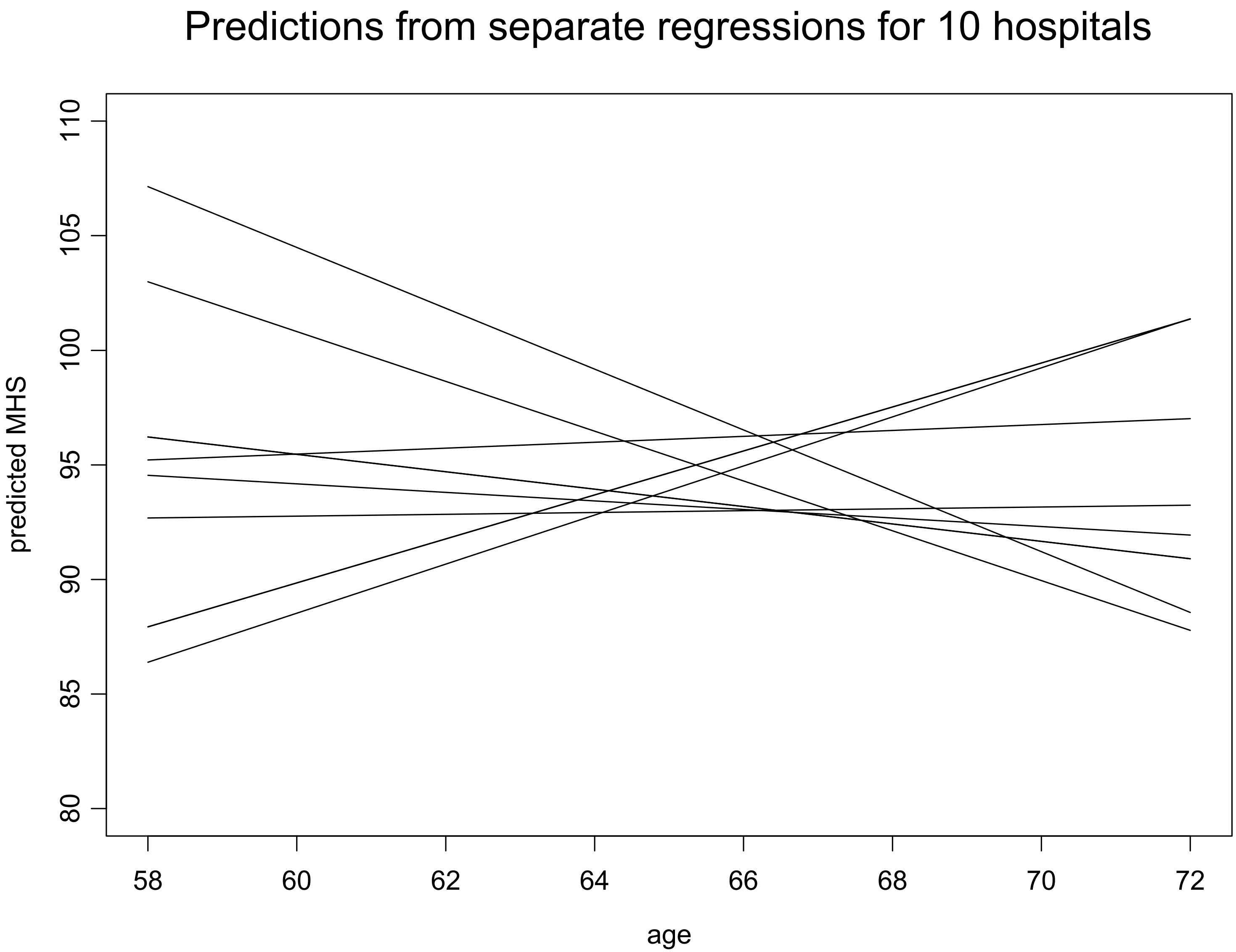
Correlation between MHS & Age = -.03 (nothing to write home about)



First Glance
Variation in MHS across Hospitals?



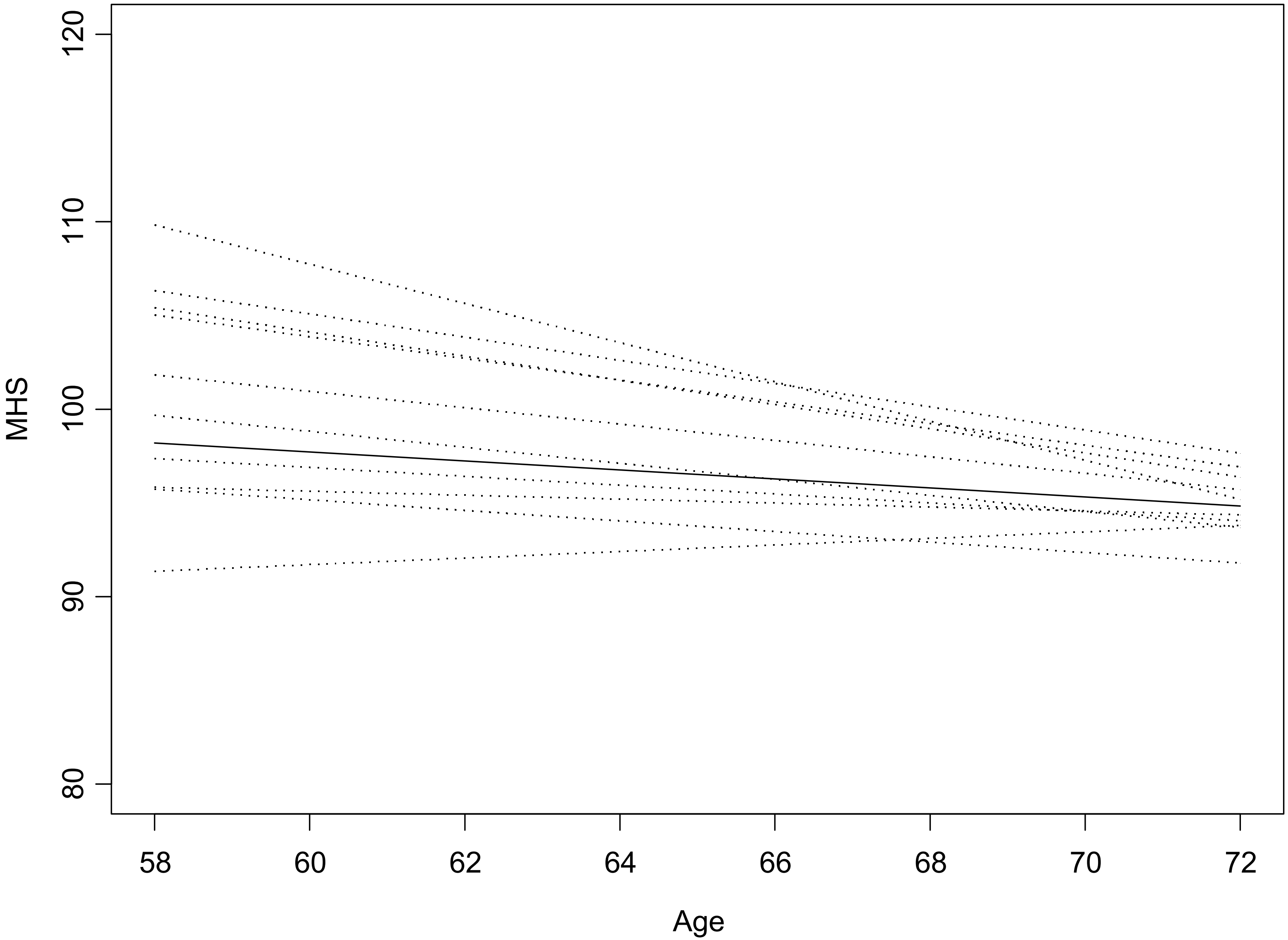
First Glance
Important Variation in effect of AGE across Hospitals?



Group Discussion: Pages 9 - 12

Example 2		Unconditional Page 9	Conditional Page 10	Random Intercepts Page 11	Random Intercepts & Slopes Page 12
Fixed Effects (Level 1)	Intercept (age 65)	95.3	95.3	96.5	96.5
	Centered Age		-.13 (10)	-.12 (.10)	-.23 (.11)
Level 1	Residual	108.5	108.5	99.2	98.3
Random Effects (Level 2)	(1,1) Intercepts			13.8	13.7
	(1,2)				-1.3
	(2,2) Slopes				.24
	AIC	19303	19304	19215	19211
	-2LL	19301	19302	19211	19203

Example #2, Model from Page 13, Age Fixed Effect, Random Slope & Intercept for Hospitals



Other things to consider

- Variance Structure
  - We used “unstructured” but there are MANY others
  - Be careful; some software uses same term for different variance
- Estimation Method
  - We used “REML”
  - REML method will be VERY slow (and might not work) on large data sets

3 Things to Remember about Fitting Data to Multilevel Models

1. Look at your data – make eye contact
   - Check the # of clusters
   - Check the range of sample size across clusters
   - Be creative - Graph something!
2. Center independent variables; let the reference group represent a typical person
   - Important for interpretation
   - Important for convergence (be nice to your software!)
3. Model fitting is ITERATIVE – particularly for Multilevel Models
   - you will have to fit more than 1 model before you are finished!

Sources of Information that might be helpful when you use multilevel models (Page 13)

1. Carolina Population Center, *A SAS User's Guide to Stata*
http://www.cpc.unc.edu/services/computer/presentations/sas_to_stata/sas_to_stata.html
This site does not include information for running "mixed" SAS models in Stata but it is a great general guide to Stata if you are used to programming in SAS.

2. UCLA Academic Technology Services
http://www.ats.ucla.edu/stat/
A great site for many needs!
   a. http://www.ats.ucla.edu/stat/sas/default.htm specifics for SAS
   b. http://www.ats.ucla.edu/stat/stata/default.htm specifics for Stata
   c. http://www.ats.ucla.edu/stat/examples/alda.htm Gives code for running data examples in Singer & Willet *Applied Longitudinal Data*. Extra nice because it gives code in Mplus, MLwiN, HLM, SAS, Stata, R (S+) and some SPSS!

3. Singer, J. D. (1998). Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models. *Journal of Educational and Behavioral Statistics*, Vol. 24, 323-355.

4. SAS online doc 9.1 about the mixed procedure http://support.sas.com/91doc/docMainpage.jsp

5. Centre for Multilevel Modelling, Research Unit, University of Bristol; distributor for MLwiN software. http://www.cmm.bristol.ac.uk/

Thank you for attending and for your participation!

- Cindy: cindylc@bu.edu
- Susan: slvland@bu.edu

CHQOER
Bedford Center for Health Quality, Outcomes, & Economics Research

BOSTON UNIVERSITY

MODEL WITH 3 LEVELS

FROM http://ssc.utexas.edu/consulting/answers/sas/sas99.html

```
PROC MIXED DATA = sasdata  METHOD = ML  COVTEST    ;
    CLASS school classrm ;
    MODEL math = english / SOLUTION ;
    RANDOM INT / TYPE = UN SUBJECT = school ;
    RANDOM INT / TYPE = UN SUBJECT = classrm(school) ;
    TITLE 'Three-level Junior School Project model';
  RUN ;
```

The **PROC MIXED** statement lists the SAS dataset used in the analysis, **sasdata**. It also uses the **ML** or maximum likelihood estimation method. The **COVTEST** option requests covariance parameter estimates and associated test statistics be printed on the output.

The **CLASS** statement tells **PROC MIXED** that **school** and **classrm** are classification variables. The **MODEL** statement tells **PROC MIXED** that the dependent variable, **math**, is a function of the intercept or grand mean estimate and the English test score variable. Like all SAS regression and general linear model procedures, **PROC MIXED** assumes the presence of an intercept in the model unless the user explicitly specifies an option (**NOINT**) telling **PROC MIXED** to fit a no intercept model. The **SOLUTION** option has **PROC MIXED** print out regression parameter estimates, standard error estimates, and associated test statistics in tabular form.

There are two **RANDOM** statements shown in the **PROC MIXED** syntax. The first **RANDOM** statement features the keyword **INT**, which tells **PROC MIXED** to estimate separate intercept values for each classroom and school. The **TYPE = UN** option tells **PROC MIXED** to use an unstructured covariance matrix for the random effects and the **SUBJECT = school** option tells **PROC MIXED** that the clustering variable is the school. You may also include a **SOLUTION** option on the the **RANDOM** statement to obtain parameter estimates for the individual classrooms and schools.

The second **RANDOM** statement is identical to the first, except that instead of using school as the clustering variable as is the case in the first random statement we now use **classrm(school)** as the clustering variable. SAS interprets **classrm(school)** as "classroom within school". With the inclusion of both **RANDOM** statements the **PROC MIXED** syntax now estimates variances for intercepts at the school level *and* the classroom within school level, as well as the covariances between these random parameter estimates. These statistics represent the amount of variance attributable to school and classroom membership, and the relationships between schools and classrooms intercepts.

CONDITIONAL MODEL: FIXED EFFECT OF AGE BUT NO RANDOM EFFECTS

```
PROC mixed DATA=in.cindy_data2 covtest;
title2 'MIXED covtest - MODEL mh_score_sameslope=age/solution';
MODEL mh_score_sameslope=age/solution;
RUN; quit;
```

Model Information

Data Set	IN.CINDY_DATA2
Dependent Variable	mh_score_sameslope
Covariance Structure	Diagonal
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Residual

Dimensions

Covariance Parameters	1
Columns in X	2
Columns in Z	0
Subjects	1
Max Obs Per Subject	2565

Number of Observations

Number of Observations Read	2565
Number of Observations Used	2565
Number of Observations Not Used	0

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z
Residual	123.83	3.4592	35.80	<.0001

Fit Statistics

-2 Res Log Likelihood	19641.5
AIC (smaller is better)	19643.5
AICC (smaller is better)	19643.5
BIC (smaller is better)	19649.4

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	81.2689	6.9667	2563	11.67	<.0001
age	0.2866	0.1072	2563	2.67	0.0076

PROC glm gives the same answer as PROC mixed

```
PROC glm DATA=in.cindy_data2;
title2 'GLM - MODEL mh_score_sameslope=age/solution';
MODEL mh_score_sameslope=age/solution;
RUN; quit;
```

```
The GLM Procedure

Number of Observations Read        2565
Number of Observations Used        2565

The GLM Procedure

Dependent Variable: mh_score_sameslope

                                        Sum of
Source                     DF          Squares     Mean Square    F Value
Pr > F

Model                       1         884.9866        884.9866       7.15
0.0076

Error                    2563      317379.2596        123.8312

Corrected Total          2564      318264.2462


R-Square     Coeff Var      Root MSE    mh_score_sameslope Mean

0.002781      11.14087      11.12794                   99.88402


                                  Standard
Parameter          Estimate          Error    t Value    Pr > |t|

Intercept       81.26893187     6.96671265      11.67      <.0001
age              0.28656649     0.10719438       2.67      0.0076
```

UNCONDITIONAL MODEL – TOTAL VARIANCE WITHOUT FIXED AND WITHOUT RANDOM EFFECTS

```
PROC mixed DATA=in.cindy_data2 covtest;
title2 'MIXED covtest - MODEL mh_score_sameslope= /solution';
MODEL mh_score_sameslope= /solution;
RUN; quit;
```

```
The Mixed Procedure

                  Model Information

Data Set                      IN.CINDY_DATA2
Dependent Variable            mh_score_sameslope
Covariance Structure          Diagonal
Estimation Method             REML
Residual Variance Method      Profile
Fixed Effects SE Method       Model-Based
Degrees of Freedom Method     Residual


            Dimensions

Covariance Parameters             1
Columns in X                      1
Columns in Z                      0
Subjects                          1
Max Obs Per Subject            2565


          Number of Observations

Number of Observations Read             2565
Number of Observations Used             2565
Number of Observations Not Used            0

            Covariance Parameter Estimates

                         Standard         Z
Cov Parm     Estimate       Error     Value        Pr Z

Residual       124.13      3.4668     35.81      <.0001

           Fit Statistics

-2 Res Log Likelihood         19646.0
AIC (smaller is better)       19648.0
AICC (smaller is better)      19648.0
BIC (smaller is better)       19653.9

                       Solution for Fixed Effects

                                  Standard
Effect                Estimate       Error      DF    t Value     Pr > |t|

Intercept              99.8840      0.2200    2564     454.05       <.0001
```

RANDOM INTERCEPTS FOR HOSPITALS AND AGE FIXED EFFECTS

```
PROC MIXED DATA=in.cindy_data2 covtest;
title2 'MIXED covtest; class sta; MODEL mh_score_sameslope=age/solution;
RANDOM int / subject=sta ;';
CLASS sta;
MODEL mh_score_sameslope=age/solution ddfm=bw;
RANDOM int / subject=sta type=un;
RUN; quit;
```

```
                  Model Information – SAME AS PREVIOUS

             Class Level Information – SAME AS PREVIOUS

            Dimensions

Covariance Parameters            2
Columns in X                     2
Columns in Z Per Subject         1
Subjects                       100
Max Obs Per Subject             50


          Number of Observations – SAME AS PREVIOUS

                   Convergence criteria met.

                  Covariance Parameter Estimates
                                    Standard         Z
Cov Parm     Subject    Estimate       Error     Value        Pr Z

UN(1,1)      sta         22.5315      4.0028      5.63      <.0001
Residual                  101.62      2.8928     35.13      <.0001

           Fit Statistics
-2 Res Log Likelihood        19307.4
AIC (smaller is better)      19311.4
AICC (smaller is better)     19311.4
BIC (smaller is better)      19316.6

                       Solution for Fixed Effects

                                  Standard
Effect                Estimate       Error      DF    t Value    Pr > |t|

Intercept              78.7023      6.4190      99      12.26      <.0001
age                     0.3222     0.09846    2464       3.27      0.0011
```

TRY RANDOM INTERCEPT AND SLOPES WITH AGE FIXED EFFECTS

```
PROC MIXED DATA=in.cindy_data2 covtest;
title2 'MIXED covtest; class sta; MODEL mh_score_sameslope=age/solution;
RANDOM int age/ sub=sta ;';
CLASS sta;
MODEL mh_score_sameslope=centered_age/solution ddfm=bw;
RANDOM int AGE/ subject=sta type=un;
RUN; quit;
```

```
                  Model Information – SAME AS PREVIOUS

             Class Level Information – SAME AS PREVIOUS

            Dimensions

Covariance Parameters            4
Columns in X                     2
Columns in Z Per Subject         2
Subjects                       100
Max Obs Per Subject             50


          Number of Observations – SAME AS PREVIOUS

                     Iteration History

Iteration    Evaluations    -2 Res Log Like       Criterion

        0              1     19641.50320947
        1              2     20371.61345768      0.00820025
        2              1     20316.71210782      0.02485562
        3              1     20261.43913630      0.07610033
…
       44              1     19426.26183053    22914568.715
       45              3     19423.61289844       .
       46              1     19420.66092771       .
       47              1     19417.73292534       .

WHOOPS!!!!!!!
                   WARNING: Did not converge.


  Covariance Parameter Values
       At Last Iteration

Cov Parm     Subject    Estimate

UN(1,1)      sta               0
UN(2,1)      sta         -7.1270
UN(2,2)      sta          0.2559
Residual                  101.18
```

IT IS IMPORTANT TO CENTER INDEPENDENT VARIABLES – LET'S START OVER WITH FIXED EFFECT FOR AGE AND NO RANDOM EFFECTS (CONDITIONAL MODEL)

```
PROC mixed DATA=in.cindy_data2 covtest;
title2 'MIXED covtest - MODEL mh_score_sameslope=centered_age/solution';
MODEL mh_score_sameslope=centered_age/solution;
RUN; quit;
```

```
                  Model Information

Data Set                     IN.CINDY_DATA2
Dependent Variable           mh_score_sameslope
Covariance Structure         Diagonal
Estimation Method            REML
Residual Variance Method     Profile
Fixed Effects SE Method      Model-Based
Degrees of Freedom Method    Residual

            Dimensions

Covariance Parameters             1
Columns in X                      2
Columns in Z                      0
Subjects                          1
Max Obs Per Subject            2565


          Number of Observations

Number of Observations Read             2565
Number of Observations Used             2565
Number of Observations Not Used            0


            Covariance Parameter Estimates

                         Standard         Z
Cov Parm     Estimate       Error     Value        Pr Z

Residual       123.83      3.4592     35.80      <.0001


           Fit Statistics

-2 Res Log Likelihood         19641.5
AIC (smaller is better)       19643.5
AICC (smaller is better)      19643.5
BIC (smaller is better)       19649.4

                       Solution for Fixed Effects
                                  Standard
Effect                Estimate       Error      DF    t Value    Pr > |t|

Intercept              99.8958      0.2198    2563     454.56      <.0001
centered_age            0.2866      0.1072    2563       2.67      0.0076
```

RANDOM INTERCEPTS FOR HOSPITALS AND FIXED AGE EFFECTS WITH CENTERED-AGE

```
PROC MIXED DATA=in.cindy_data2 covtest;
title2 'MIXED covtest; class sta; MODEL
mh_score_sameslope=centered_age/solution;RANDOM int/ subj=sta;';
CLASS sta;
MODEL mh_score_sameslope=centered_age/solution ddfm=bw;
RANDOM int / subject=sta type=un;
RUN; quit;
```

Model Information – SAME AS PREVIOUS

Class Level Information – SAME AS PREVIOUS

Dimensions

Covariance Parameters	2
Columns in X	2
Columns in Z Per Subject	1
Subjects	100
Max Obs Per Subject	50

Number of Observations – SAME AS PREVIOUS

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	sta	22.5315	4.0028	5.63	<.0001
Residual		101.62	2.8928	35.13	<.0001

Fit Statistics

-2 Res Log Likelihood	19307.4
AIC (smaller is better)	19311.4
AICC (smaller is better)	19311.4
BIC (smaller is better)	19316.6

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	99.6457	0.5363	99	185.80	<.0001
centered_age	0.3222	0.09846	2464	3.27	0.0011

RANDOM SLOPES AND INTERCEPTS FOR HOSPITALS WITH CENTERED-AGE FIXED EFFECT

```
PROC MIXED DATA=in.cindy_data2 covtest;
title2 'MIXED covtest; class sta; MODEL
mh_score_sameslope=centered_age/solution;RANDOM int centered_age/ subj=sta;';
CLASS sta;
MODEL mh_score_sameslope=centered_age/solution ddfm=bw;
RANDOM int centered_age / subject=sta type=un;
RUN; quit;
```

Model Information – SAME AS PREVIOUS

Class Level Information – SAME AS PREVIOUS

Dimensions

Covariance Parameters	4
Columns in X	2
Columns in Z Per Subject	2
Subjects	100
Max Obs Per Subject	50

Number of Observations – SAME AS PREVIOUS

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	sta	22.5389	4.0041	5.63	<.0001
UN(2,1)	sta	0.1540	0.5776	0.27	0.7897
UN(2,2)	sta	1.85E-17	.	.	.
Residual		101.63	2.8928	35.13	<.0001

Fit Statistics

-2 Res Log Likelihood	19307.3
AIC (smaller is better)	19313.3
AICC (smaller is better)	19313.3
BIC (smaller is better)	19321.1

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	99.6468	0.5363	99	185.79	<.0001
centered_age	0.3205	0.09843	2464	3.26	0.0011

START WITH THE UNCONDITIONAL MODEL

```
PROC mixed DATA=in.cindy_data4 covtest;
title2 'MIXED covtest - MODEL mh_score_diffslope= /solution';
MODEL mh_score_diffslope= /solution;
RUN; quit;
```

```
                  Model Information

Data Set                     IN.CINDY_DATA4
Dependent Variable           mh_score_diffslope
Covariance Structure         Diagonal
Estimation Method            REML
Residual Variance Method     Profile
Fixed Effects SE Method      Model-Based
Degrees of Freedom Method    Residual

            Dimensions

Covariance Parameters            1
Columns in X                     1
Columns in Z                     0
Subjects                         1
Max Obs Per Subject           2565


          Number of Observations

Number of Observations Read           2565
Number of Observations Used           2565
Number of Observations Not Used          0


            Covariance Parameter Estimates

                         Standard         Z
Cov Parm     Estimate       Error     Value        Pr Z

Residual       108.51      3.0305     35.81      <.0001

           Fit Statistics

-2 Res Log Likelihood        19301.1
AIC (smaller is better)      19303.1
AICC (smaller is better)     19303.1
BIC (smaller is better)      19309.0


                       Solution for Fixed Effects

                                   Standard
Effect                 Estimate       Error       DF    t Value     Pr > |t|

Intercept               95.3253      0.2057     2564     463.47       <.0001
```

TRY THE CONDITIONAL MODEL WITH CENTERED-AGE FIXED EFFECT

```
PROC mixed DATA=in.cindy_data4 covtest;
title2 'MIXED covtest - MODEL mh_score_diffslope=centered_age/solution';
MODEL mh_score_diffslope=centered_age/solution;
RUN; quit;
```

```
MIXED covtest - MODEL mh_score_diffslope=centered_age/solution

The Mixed Procedure

                  Model Information

Data Set                     IN.CINDY_DATA4
Dependent Variable           mh_score_diffslope
Covariance Structure         Diagonal
Estimation Method            REML
Residual Variance Method     Profile
Fixed Effects SE Method      Model-Based
Degrees of Freedom Method    Residual

            Dimensions

Covariance Parameters             1
Columns in X                      2
Columns in Z                      0
Subjects                          1
Max Obs Per Subject            2565

            Covariance Parameter Estimates

                         Standard         Z
Cov Parm      Estimate      Error     Value        Pr Z

Residual        108.48     3.0303     35.80      <.0001

           Fit Statistics

-2 Res Log Likelihood         19302.3
AIC (smaller is better)       19304.3
AICC (smaller is better)      19304.3
BIC (smaller is better)       19310.1

                       Solution for Fixed Effects

                                   Standard
Effect                  Estimate      Error      DF    t Value    Pr > |t|

Intercept                95.3201     0.2057    2563     463.41      <.0001
centered_age             -0.1285     0.1003    2563      -1.28      0.2006
```

ADD RANDOM INTERCEPTS FOR HOSPITALS

```
PROC MIXED DATA=in.cindy_data4 covtest;
title2 'MIXED covtest; class sta; MODEL mh_score_diffslope=centered_age/s;
RANDOM int / sub=sta ;';
CLASS sta;
MODEL mh_score_diffslope=centered_age/solution ddfm=bw;
RANDOM int / subject=sta type=un;
RUN; quit;
```

```
                  Model Information  - SAME as PREVIOUS

             Class Level Information – SAME as PREVIOUS

            Dimensions

Covariance Parameters             2
Columns in X                      2
Columns in Z Per Subject          1
Subjects                        100
Max Obs Per Subject              50

Number of Observations – SAME as PREVIOUS

                   Convergence criteria met.

                  Covariance Parameter Estimates
```

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	sta	13.7747	2.9989	4.59	<.0001
Residual		99.2002	2.8361	34.98	<.0001

Fit Statistics

-2 Res Log Likelihood	19210.9
AIC (smaller is better)	19214.9
AICC (smaller is better)	19214.9
BIC (smaller is better)	19220.1

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	96.5423	0.4405	99	219.19	<.0001
centered_age	-0.1181	0.09715	2464	-1.22	0.2242

ADD RANDOM SLOPES FOR HOSPITALS

```
PROC MIXED DATA=in.cindy_data4 covtest;
title2 'MIXED covtest; class sta; MODEL
mh_score_diffslope=centered_age/s;RANDOM int centered_age/ sub=sta;';
CLASS sta;
MODEL mh_score_diffslope=centered_age/solution ddfm=bw;
RANDOM int centered_age  / subject=sta type=un;
RUN; quit;
```

```
                  Model Information  - SAME as PREVIOUS

             Class Level Information – SAME as PREVIOUS

            Dimensions

Covariance Parameters            4
Columns in X                     2
Columns in Z Per Subject         2
Subjects                       100
Max Obs Per Subject             50


          Number of Observations – SAME as PREVIOUS

                   Convergence criteria met.

                  Covariance Parameter Estimates
```

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	sta	13.6631	2.9943	4.56	<.0001
UN(2,1)	sta	-1.2676	0.5829	-2.17	0.0297
UN(2,2)	sta	0.2425	0.1647	1.47	0.0704
Residual		98.3452	2.8578	34.41	<.0001

```
           Fit Statistics

-2 Res Log Likelihood         19203.4
AIC (smaller is better)       19211.4
AICC (smaller is better)      19211.4
BIC (smaller is better)       19221.8

                       Solution for Fixed Effects
```

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	96.5216	0.4389	99	219.90	<.0001
centered_age	-0.2264	0.1111	2464	-2.04	0.0416

Sources of Information that might be helpful when you use multilevel models:

1. Carolina Population Center, *A SAS User's Guide to Stata*
http://www.cpc.unc.edu/services/computer/presentations/sas_to_stata/sas_to_stata.html
This site does not include information for running "mixed" SAS models in Stata but it is a great general guide to Stata if you are used to programming in SAS.

2. UCLA Academic Technology Services
http://www.ats.ucla.edu/stat/
A great site for many needs!
    a. http://www.ats.ucla.edu/stat/sas/default.htm specifics for SAS
    b. http://www.ats.ucla.edu/stat/stata/default.htm specifics for Stata
    c. http://www.ats.ucla.edu/stat/examples/alda.htm Gives code for running data examples in Singer & Willet *Applied Longitudinal Data*. Extra nice because it gives code in Mplus, MLwiN, HLM, SAS, Stata, R (S+) and some SPSS!

3. Singer, J. D. (1998). Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models. *Journal of Educational and Behavioral Statistics*, Vol. 24, 323-355.

4. SAS online doc 9.1 about the mixed procedurehttp://support.sas.com/91doc/docMainpage.jsp

5. Centre for Multilevel Modelling, Research Unit, University of Bristol; distributor for MLwiN software. http://www.cmm.bristol.ac.uk/