

# **Missing Data in Longitudinal Studies: Mixed-effects Pattern-Mixture and Selection Models**

Hedeker D & Gibbons RD (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*, 2, 64-78.

Chapter 14 in Hedeker & Gibbons (2006), *Longitudinal Data Analysis*, Wiley.

## Example: Treatment-Related Change Across Time

Data from the NIMH Schizophrenia collaborative study on treatment related changes in overall severity. IMPS item 79, *Severity of Illness*, was scored as:

- 1 = normal
- 2 = borderline mentally ill
- 3 = mildly ill
- 4 = moderately ill
- 5 = markedly ill
- 6 = severely ill
- 7 = among the most extremely ill

---

---

Group	Sample size at Week							<i>completers</i>
	0	1	2	3	4	5	6	
PLC (n=108)	107	105	5	87	2	2	70	65%
DRUG (n=329)	327	321	9	287	9	7	265	81%

---

---

*Drug = Chlorpromazine, Fluphenazine, or Thioridazine*

## Descriptive Statistics - NIMH Schizophrenia study

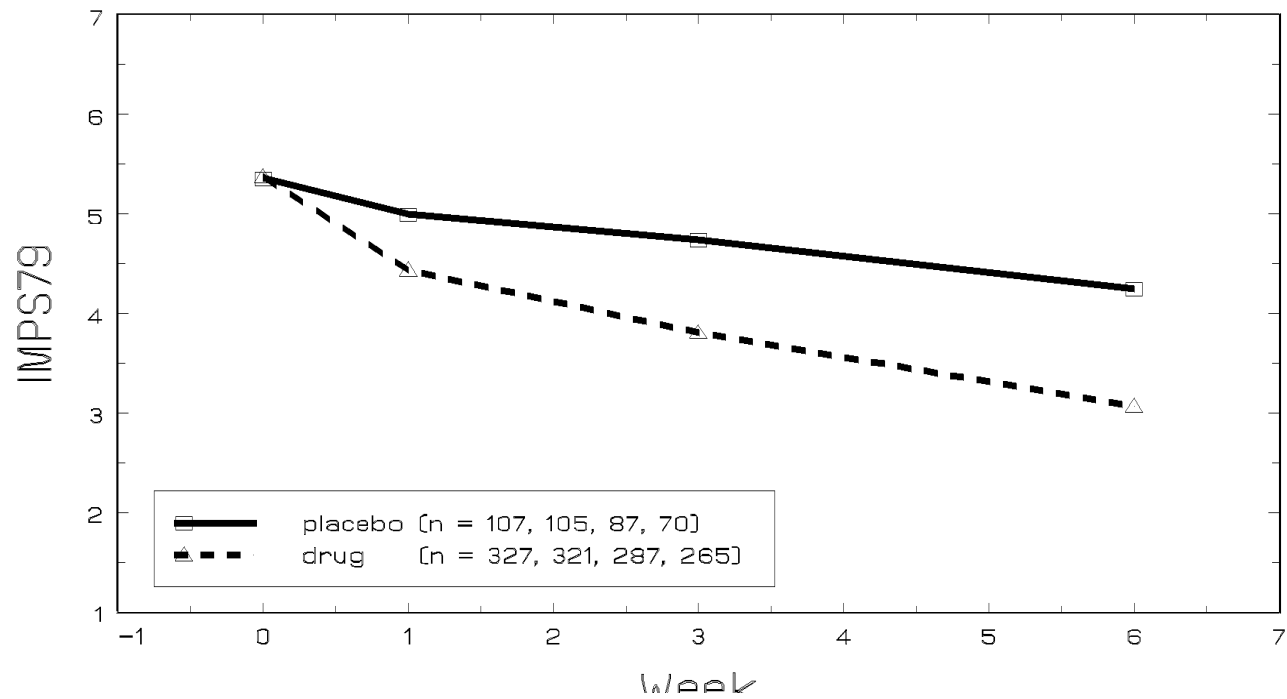
Observed IMPS79 Means,  $n$ , and sd

	<u>week 0</u>	<u>week 1</u>	<u>week 3</u>	<u>week 6</u>
placebo	5.35	4.99	4.74	4.25
$n$	107	105	87	70
drug	5.37	4.43	3.80	3.06
$n$	327	321	287	265
pooled sd	.87	1.23	1.44	1.48

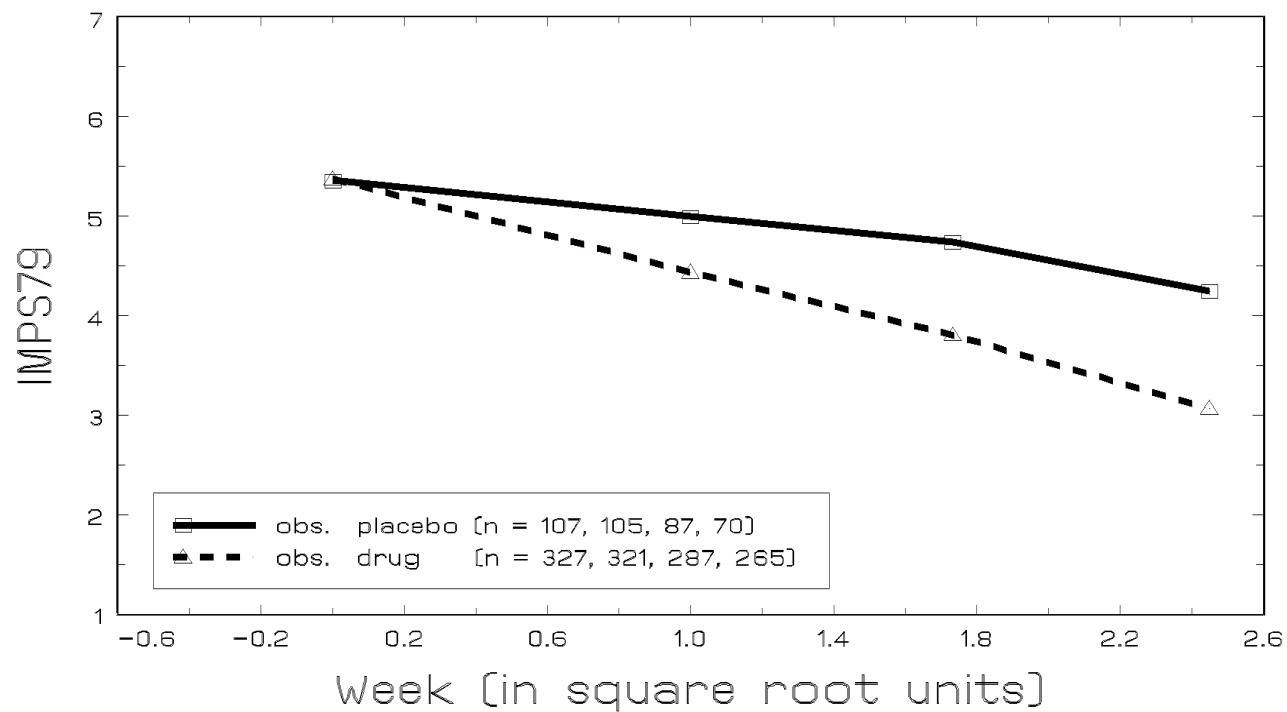
Correlations:  $n = 313$  and  $321 \leq n \leq 424$

	<u>week 0</u>	<u>week 1</u>	<u>week 3</u>	<u>week 6</u>
week 0	1.0	.41	.25	.15
week 1	<b>.43</b>	1.0	.67	.47
week 3	<b>.29</b>	<b>.67</b>	1.0	.67
week 6	<b>.14</b>	<b>.47</b>	<b>.68</b>	1.0

Mean IMPS79 across Time by Group



Mean IMPS79 across Time by Group



## Mixed-effects regression model (MRM) - Schizophrenia study

$$IMPS79 = Drug + Time + (Drug \times Time) + Subj + (Subj \times Time) + Error$$

$$IMPS79_{ij} = \beta_0 + \beta_1 Drug_i + \beta_2 SWeek_j + \beta_3 (Drug_i \times SWeek_j) + v_{0i} + v_{1i} SWeek_j + \varepsilon_{ij}$$

$$i = 1, \dots, N \text{ subjects} \quad j = 1, \dots, n_i \text{ obs}$$

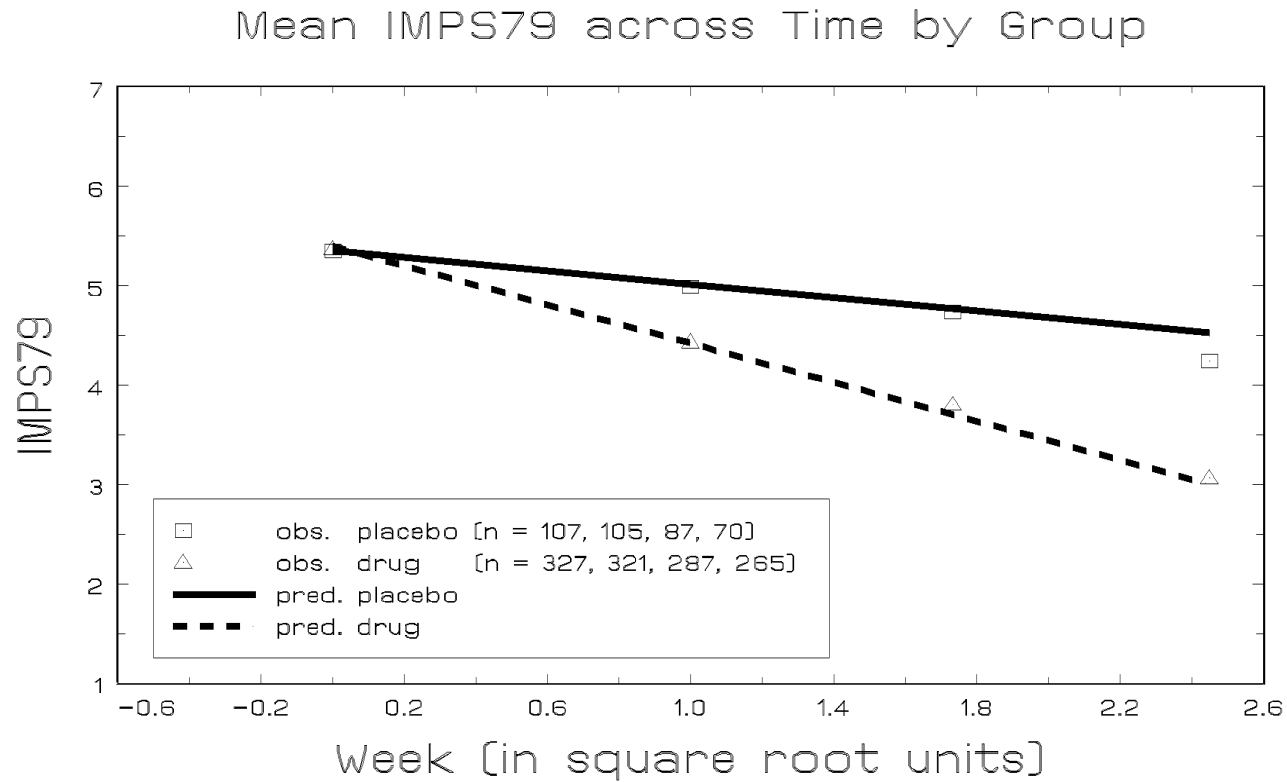
Drug = 0 for placebo, 1 for drug

SWeek = 0,  $\sqrt{1} = 1$ ,  $\sqrt{2} = 1.41$ ,  $\sqrt{3} = 1.73$ ,  $\sqrt{4} = 2$ ,  $\sqrt{5} = 2.24$ ,  $\sqrt{6} = 2.45$

NIMH Schizophrenia Study - IMPS79 across Time: ML Estimates (se)

	<i>Completers</i>			<i>All Subjects</i>		
	<i>N = 335</i>			<i>N = 437</i>		
	est.	se	<i>p</i> <	est.	se	<i>p</i> <
intercept	5.221	0.109	.001	5.348	0.088	.001
Drug (0 = plc; 1 = drug)	0.202	0.123	.101	0.046	0.101	.65
Time (sqrt week)	-0.393	0.073	.001	-0.336	0.068	.001
Drug by Time	-0.539	0.083	.001	-0.641	0.078	.001
Intercept variance	0.398	0.068		0.369	0.060	
	<i>sd = .63</i>			<i>sd = .61</i>		
Int-Time covariance	-0.011	0.035		0.021	0.034	
	<i>r = -.04</i>			<i>r = .07</i>		
Time variance	0.205	0.031		0.242	0.032	
	<i>sd = .45</i>			<i>sd = .49</i>		

# Fitted and Obs. Means across Time by Condition



$$\hat{IMPS}_{ij} = 5.35 + .05 Drug_i - .34 Time_j - .64(D_i \times T_j)$$

## Missing Data and Incomplete Data Models

$$\mathbf{y} \begin{cases} \mathbf{y}^{(O)} & R = 0 \text{ (observed)} \\ \mathbf{y}^{(M)} & R = 1 \text{ (missing)} \end{cases}$$

- GEE assumes “Missing Completely at Random” (MCAR)

$$P(\mathbf{R} \mid \mathbf{y}, \mathbf{X}) = P(\mathbf{R} \mid \mathbf{X}) \text{ for all } \mathbf{y}$$

conditional on covariates,  $\mathbf{R}$  is independent of  $\mathbf{y}^{(O)}$  and  $\mathbf{y}^{(M)}$

$\Rightarrow$  “covariate-dependent missingness”

- Likelihood-based MRM assumes “Missing at Random” (MAR)

$$P(\mathbf{R} \mid \mathbf{y}, \mathbf{X}) = P(\mathbf{R} \mid \mathbf{X}, \mathbf{y}^{(O)}) \text{ for all } \mathbf{y}^{(M)}$$

conditional on covariates *and* observed values of the dependent variable,  $\mathbf{R}$  is independent of  $\mathbf{y}^{(M)}$

$\Rightarrow$  “ignorable non-response”

## Missing Not At Random (MNAR) Models

- When the data are nonignorable (*i.e.*, MNAR), standard statistical models can yield badly biased results
- The observed data provide no information to either confirm or refute ignorability

⇒ **cannot test MAR versus MNAR**

Two general classes of MNAR models

- Pattern mixture models - use missing data pattern information in the longitudinal modeling
- Selection models - modeling of both the longitudinal and missingness processes

⇒ will be illustrated in terms of MRMs, however they can be more broadly defined and utilized

## Comments on MNAR models

- Ordinary MRM (and other full-likelihood models) assume MAR, these extended models do not
- Use of nonignorable models can be helpful in conducting a sensitivity analysis; to see how the conclusions might vary as a function of what is assumed about the missing data
- Not necessarily a good idea to rely on a single MNAR model, because the assumptions about the missing data are impossible to assess with the observed data
- One should use MNAR models sensibly, possibly examining several types of such models for a given dataset

## Pattern-mixture models for missing data

*Little (1993, 1994, 1995); Hedeker & Gibbons (1997)*

- divide subjects into groups depending on their missing data pattern
- the missing data pattern is a between-subjects variable to be used in longitudinal data analysis

For 3 timepoints, there are eight ( $2^3$ ) possible missing data patterns:

pattern group	time1	time2	time3
1	O	O	O
2	O	O	M
3	O	M	O
4	M	O	O
5	M	M	O
6	O	M	M
7	M	O	M
8	M	M	M

where, O=observed and M=missing

Since MMM provides no data, it is ignored in the analysis

## *Representing patterns with dummy-coded variables*

pattern	$D1$	$D2$	$D3$	$D4$	$D5$	$D6$
OOO	0	0	0	0	0	0
OOM	1	0	0	0	0	0
OMO	0	1	0	0	0	0
MOO	0	0	1	0	0	0
MMO	0	0	0	1	0	0
OMM	0	0	0	0	1	0
MOM	0	0	0	0	0	1

- these dummy-codes represent deviations from pattern OOO
- Other coding schemes can be used (“effect” or “sequential” coding)
- these variables are used as main effects and interactions

## Classification of Subjects based on missing-data

$$\text{Drop}_i = \begin{cases} 0 & \text{subject measured at week 6 (last timepoint)} \\ 1 & \text{subject missing at week 6 (last timepoint)} \end{cases}$$

Drug group	Drop group		total
	completer	dropout	
placebo	70 (.65)	38 (.35)	108
drug	265 (.81)	64 (.19)	329
total	335	102	437

- Dropout not independent of Drug  $\chi_1^2 = 11.25, p < .001$
- Is dropout related to severity of illness?
- Does dropout moderate the influence of other variables' effects on severity of illness?

## Mixed-effects pattern mixture model: Schiz data

augment the basic MRM of IMPS79 over time:

$$\text{IMPS79}_{ij} = \beta_0 + \beta_1 \text{Drug}_i + \beta_2 \text{SWeek}_j + \beta_3 (\text{Drug}_i \times \text{SWeek}_j) + \nu_{0i} + \nu_{1i} \text{SWeek}_j + \varepsilon_{ij} ,$$

with variables based on the missing data patterns

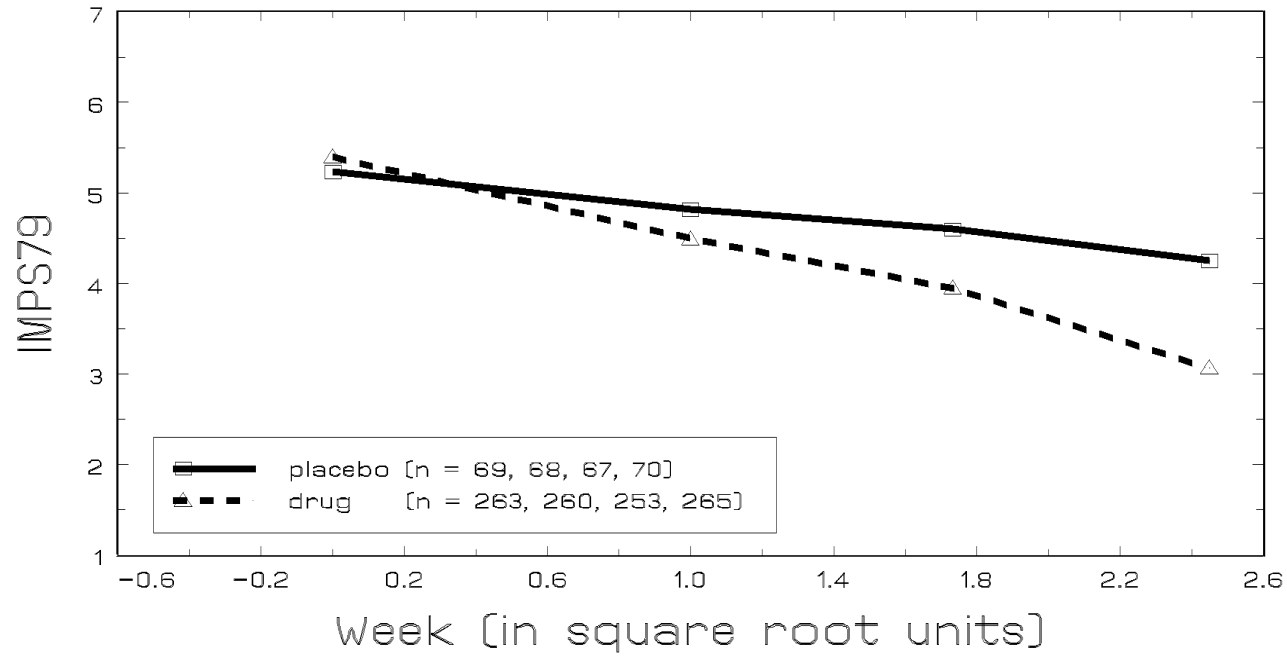
*e.g.*, completers (N = 335) vs non-completers (N = 102)

**Drop** = 0 or 1 for those that did not or did dropout from the trial (*i.e.*, were not measured at the final study timepoint)

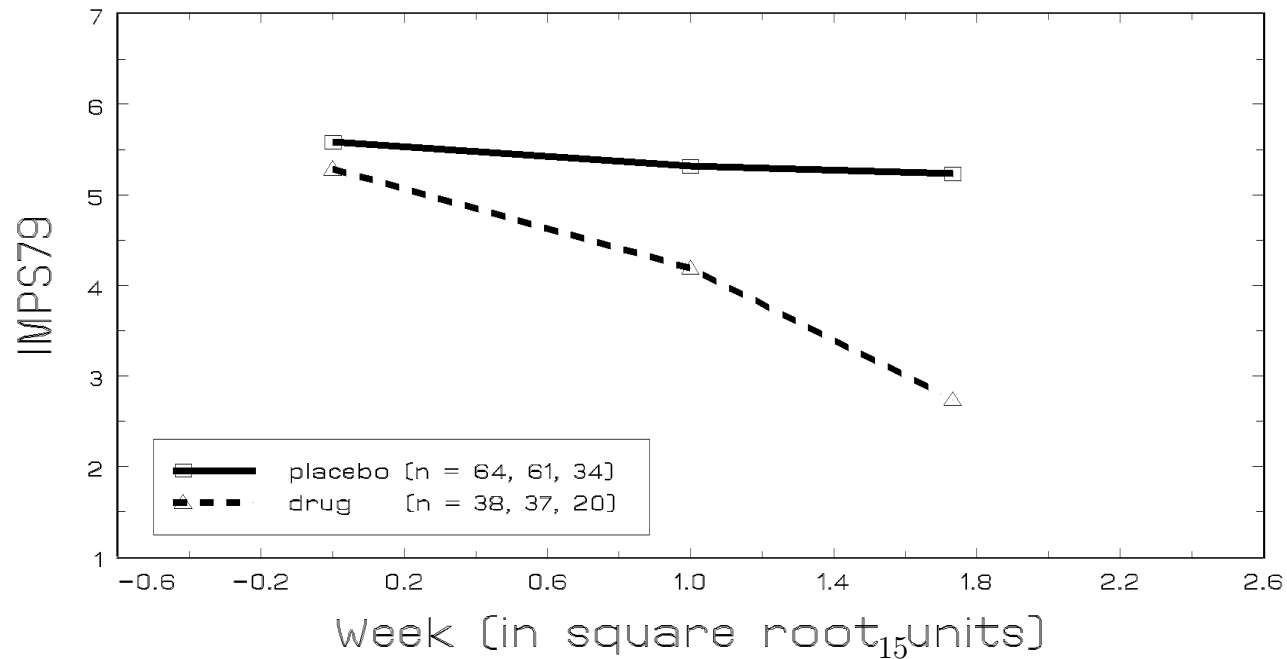
$$\begin{aligned} \text{IMPS79}_{ij} = & \beta_0 + \beta_1 \text{Drug}_i + \beta_2 \text{SWeek}_j + \beta_3 (\text{Drug}_i \times \text{SWeek}_j) \\ & + \beta_0^D \text{Drop}_i + \beta_1^D (\text{Drop}_i \times \text{Drug}_i) + \beta_2^D (\text{Drop}_i \times \text{Sweek}_j) \\ & + \beta_3^D (\text{Drop}_i \times \text{Drug}_i \times \text{Sweek}_j) + \nu_{0i} + \nu_{1i} \text{SWeek}_j + \varepsilon_{ij} \end{aligned}$$

- $\beta_0, \beta_1, \beta_2,$  and  $\beta_3$  are for the completer subsample
- $\beta_0^D, \beta_1^D, \beta_2^D,$  and  $\beta_3^D$  how dropouts differ from completers
- three-way interaction is of particular interest - indicates how the drug by time interaction varies with study completion

Mean IMPS79 across Time by Group  
Completers



Mean IMPS79 across Time by Group  
Dropouts



## Less Restrictive Pattern Mixture Model

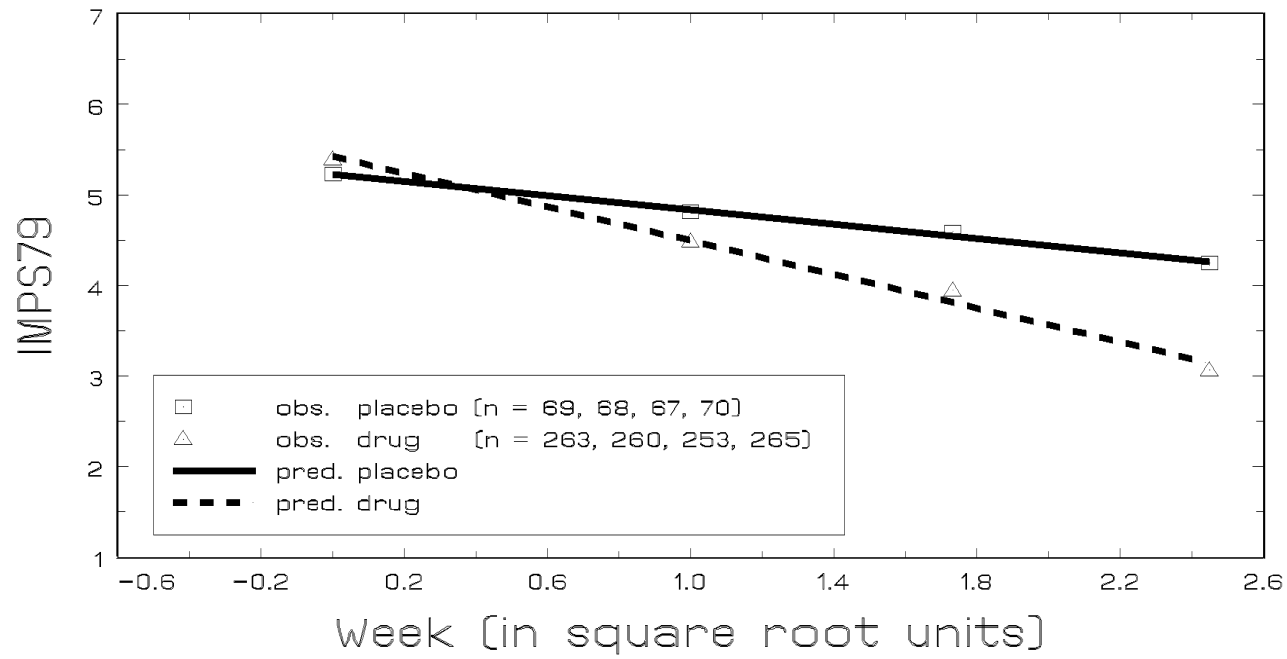
- use week of dropout variable  $D_i$  in forming missing data patterns
- six missing data patterns: five dropout weeks and completers
- Let  $D_m = D_1, \dots, D_5$  denote dummy-variables which contrast each dropout pattern to the completers

$$\begin{aligned} \text{IMPS79}_{ij} = & \beta_0 + \beta_1 \text{Drug}_i + \beta_2 \text{SWeek}_j + \beta_3 (\text{Drug}_i \times \text{SWeek}_j) \\ & + \sum_{m=1}^5 \beta_0^m D_m + \beta_1^m (D_m \times \text{Drug}_i) + \beta_2^m (D_m \times \text{Sweek}_j) \\ & + \beta_3^m (D_m \times \text{Drug}_i \times \text{Sweek}_j) \\ & + v_{0i} + v_{1i} \text{SWeek}_j + \varepsilon_{ij} \end{aligned}$$

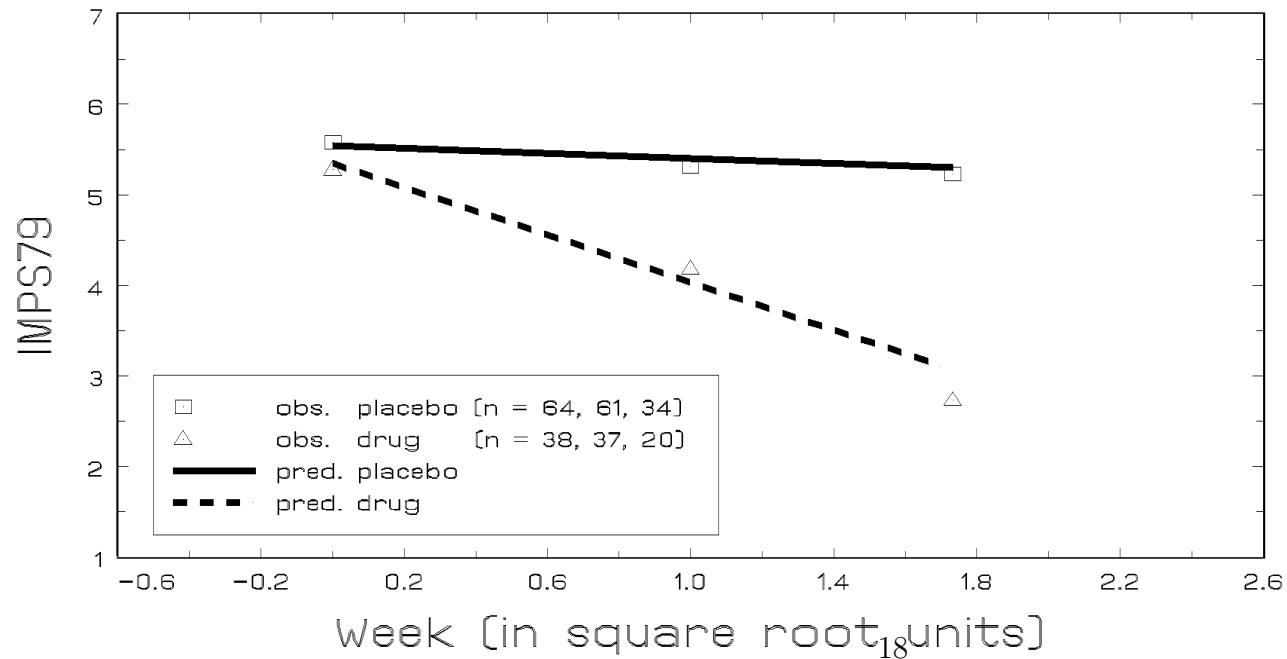
- $\beta_0, \beta_1, \beta_2, \beta_3$  are for completers
- $\beta_0^m, \beta_1^m, \beta_2^m, \beta_3^m$  indicate how dropout group  $m$  differs from completers
- the  $\beta_3^m$  parameters are of great interest

parameter	est	se	$p <$	est	se	$p <$	est	se	$p <$
Int $\beta_0$	5.348	.088	.0001	5.221	.108	.0001	5.221	.107	.0001
Drug $\beta_1$	.046	.101	.65	.202	.121	.096	.202	.120	.094
SWeek $\beta_2$	-.336	.068	.0001	-.393	.076	.0001	-.393	.075	.0001
Drug $\times$ SWeek $\beta_3$	-.641	.078	.0001	-.539	.086	.0001	-.539	.085	.0001
<u>Dropout</u>				<u>Drop = 1 (N = 102)</u>			<u>D = 1 (N = 37)</u>		
Int $\beta_0^1$				.320	.186	.086	.471	.288	.102
Drug $\beta_1^1$				-.399	.227	.079	-.456	.353	.20
SWeek $\beta_2^1$				.252	.159	.115	.240	.334	.47
Drug $\times$ SWeek $\beta_3^1$				-.635	.196	.002	-.412	.412	.32
							<u>D = 2 (N = 10)</u>		
Int $\beta_0^2$							.524	.437	.23
Drug $\beta_1^2$							-.703	.613	.25
SWeek $\beta_2^2$							.338	.398	.40
Drug $\times$ SWeek $\beta_3^2$							-.735	.562	.19
							<u>D = 3 (N = 42)</u>		
Int $\beta_0^3$							.047	.256	.85
Drug $\beta_1^3$							-.198	.318	.53
SWeek $\beta_2^3$							.377	.208	.07
Drug $\times$ SWeek $\beta_3^3$							-.835	.261	.002
							<u>D = 4 (N = 5)</u>		
Int $\beta_0^4$							.801	.653	.22
Drug $\beta_1^4$							-.237	.841	.78
SWeek $\beta_2^4$							-.101	.485	.84
Drug $\times$ SWeek $\beta_3^4$							-1.210	.625	.054
							<u>D = 5 (N = 8)</u>		
Int $\beta_0^5$							.337	.645	.60
Drug $\beta_1^5$							-.842	.746	.26
SWeek $\beta_2^5$							-.157	.466	.74
Drug $\times$ SWeek $\beta_3^5$							.231	.538	.67
Deviance	4649.0			4623.3			4607.8		

Mean IMPS79 across Time by Group  
Completers



Mean IMPS79 across Time by Group  
Dropouts



## Pattern-mixture averaged results (Little, 1995)

- Obtained averaging over missing-data patterns
  - *e.g.*, completers and dropouts
- Uses sample proportions as estimates of missing-data pattern proportions
- Depends on “model” for missing-data patterns
  - *e.g.*, completer versus dropout status varies by tx

### Completer

*placebo*      70/108

*drug*          265/329

**335/437**

### Dropout

*placebo*      38/108

*drug*          64/329

**102/437**

## Pattern-mixture averaged results

$$\hat{\beta} = \hat{\pi}_c \hat{\beta}_c + \hat{\pi}_d \hat{\beta}_d = \hat{\beta}_c + \hat{\pi}_d \hat{\beta}^D$$

- $\hat{\beta}_c$  correspond to the coefficients in the current model formulation not involving dropout (*i.e.*, intercept, drug, time, drug by time)
- $(\hat{\beta}_d - \hat{\beta}_c) = \hat{\beta}^D$  correspond to the dropout-related coefficients in the current model formulation (*i.e.*, dropout, dropout by drug, dropout by time, dropout by drug by time)
- $\hat{\pi}_d$  is the sample proportion of dropouts

⇒ averaged estimates are linear combinations of model estimates (obtained by simple arithmetic or via **ESTIMATE** statement in SAS)

## Placebo Intercept

$$\frac{335}{437}(5.22) + \frac{102}{437}(5.22 + 0.32) = 5.22 + \frac{102}{437}(0.32) = 5.30$$

*Completers*      *Dropouts*

## Placebo Time effect

$$\frac{335}{437}(-0.39) + \frac{102}{437}(-0.39 + 0.25) = -0.39 + \frac{102}{437}(0.25) = -0.33$$

*Completers*      *Dropouts*

## Drug Intercept difference

$$\frac{335}{437}(0.20) + \frac{102}{437}(0.20 - 0.40) = 0.20 + \frac{102}{437}(-0.40) = 0.11$$

*Completers*      *Dropouts*

## Drug Time difference

$$\frac{335}{437}(-0.54) + \frac{102}{437}(-0.54 - 0.64) = -0.54 + \frac{102}{437}(-0.64) = -0.69$$

*Completers*      *Dropouts*

## SAS example using ESTIMATE

```
/* pattern-mixture random intercept and trend model
/* using marginal dropout proportion to estimate averaged results */

PROC MIXED METHOD=ML COVTEST;
CLASS id;
MODEL imps79 = sweek drug sweek*drug dropout dropout*sweek
           dropout*drug dropout*drug*sweek / SOLUTION;
RANDOM INTERCEPT sweek /SUB=id TYPE=UN G GCORR;

ESTIMATE 'avg int' INTERCEPT 1 sweek 0 drug 0 sweek*drug 0 dropout .2334
           dropout*sweek 0 dropout*drug 0 dropout*drug*sweek 0;
ESTIMATE 'avg sweek' INTERCEPT 0 sweek 1 drug 0 sweek*drug 0 dropout 0
           dropout*sweek .2334 dropout*drug 0 dropout*drug*sweek 0;
ESTIMATE 'avg drug' INTERCEPT 0 sweek 0 drug 1 sweek*drug 0 dropout 0
           dropout*sweek 0 dropout*drug .2334 dropout*drug*sweek 0;
ESTIMATE 'avg sweek*drug' INTERCEPT 0 sweek 0 drug 0 sweek*drug 1 dropout 0
           dropout*sweek 0 dropout*drug 0 dropout*drug*sweek .2334;

RUN;
```

where `dropout` =0 (completers) or =1 (dropouts) and  $102/437 = .2334$  (*i.e.*, dropout proportion)

## Don't have (or like) SAS?

(*i.e.*, can I do pattern-mixture modeling with SPSS? .... YES!)

Weighted Effect Coding (Darlington, 1990, *Regression and Linear Models*, pp. 238-239): in ANOVA context

- yields comparisons of each cell mean, except the reference cell, with a weighted average of cell means

Effect coding level	$D$		Weighted effect coding level	$D^*$
1	1	becomes	1	1
2	-1		2	$-1 \times \frac{\text{weight of level 1}}{\text{weight of level 2}}$

$$Y = \beta_0 + \beta_1 D^* + e$$

$\beta_0$  = weighted average of  $Y$  means across the two levels

$\beta_1$  = difference between level 1 mean and the weighted average

# SPSS example using weighted effect coding

Step 1 - do some data management to create the necessary variables

SCHIZREP.SAV

ID	subject's ID
IMPS79	severity of illness (dependent variable)
Week	study week (= 0, 1, 2, 3, 4, 5, 6)
Drug	=0 (placebo) or =1 (drug)
SexM	=0 (female) or =1 (male)
SWeek	sqrt of Week
DrugSwk	product of Drug by SWeek
Week_max	subject's maximum value of Week
Dropout	=0 (Week_max = 6) or =1 (Week_max < 6)
DropW	$= \frac{-\hat{\pi}_D}{\hat{\pi}_C} = \frac{-102}{335}$ (Dropout = 0) or =1 (Dropout = 1)
DropWDrug	product of DropW by Drug
DropWSweek	product of DropW by SWeek
DropWDrugSwk	product of DropW by Drug by SWeek

## Model with weighted effect coding

$$\begin{aligned} \text{IMPS79}_{ij} = & \beta_0 + \beta_1 \text{Drug}_i + \beta_2 \text{SWeek}_j + \beta_3 (\text{Drug}_i \times \text{SWeek}_j) \\ & + \beta_0^{DW} \text{DropW}_i + \beta_1^{DW} (\text{DropW}_i \times \text{Drug}_i) \\ & + \beta_2^{DW} (\text{DropW}_i \times \text{Sweek}_j) \\ & + \beta_3^{DW} (\text{DropW}_i \times \text{Drug}_i \times \text{Sweek}_j) \\ & + \nu_{0i} + \nu_{1i} \text{SWeek}_j + \varepsilon_{ij} \end{aligned}$$

- $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  are weighted averages over completers and dropouts (*i.e.*, exactly what we want!)
- $\beta_0^{DW}$ ,  $\beta_1^{DW}$ ,  $\beta_2^{DW}$ , and  $\beta_3^{DW}$  how dropouts differ from the weighted averages

# SPSS syntax for model with weighted effect coding

Schizpm.sps

MIXED

Imps79 WITH Drug Sweek DrugSwk

DropW DropWDrug DropWSweek DropWDrugSwk

/FIXED = Drug Sweek DrugSwk

DropW DropWDrug DropWSweek DropWDrugSwk

/METHOD = ML

/PRINT = SOLUTION TESTCOV

/RANDOM INTERCEPT Sweek | SUBJECT(ID) COVTYPE(UN) .

## Pattern-mixture averaged results

using either approach (ESTIMATE in SAS or weighted effect coding in SPSS), we get:

parameter	estimate	std error	<i>p</i> -value
Intercept	5.2958	.0898	.0001
Drug	.1086	.1029	.29
SWeek	-.3346	.0670	.0001
DrugSwk	-.6868	.0776	.0001

Great! We're done, right?

I've got some good news, and some not-so-good news



averaged estimates are fine



standard errors are not exactly correct

## Standard Errors for Averaged Estimates

$$\hat{V}(\hat{\beta}) = \hat{V}(\hat{\beta})_F + \frac{n_d n_c}{N^3} (\hat{\beta}^D)^2$$

where,  $\hat{V}(\hat{\beta})_F$  is the variance treating the sample proportions as known, *i.e.*, the square of the standard error one gets using

$$\hat{\beta} = \hat{\beta}_c + \hat{\pi}_d \hat{\beta}^D$$

and not taking into account the fact that  $\pi_d$  is estimated (*i.e.*, this is obtained using methods that yield linear combinations of estimates and their associated standard errors)

$\Rightarrow$  simple augmentation of  $\hat{V}(\hat{\beta})_F$  to get correct standard errors

$$\text{Calculation of } \hat{V}(\hat{\beta}) = \hat{V}(\hat{\beta})_F + \frac{n_d n_c}{N^3} (\hat{\beta}^D)^2$$

parameter	$\hat{\beta}$	$\hat{V}(\hat{\beta})_F$	$\hat{\beta}^D$	Augment	$\hat{V}(\hat{\beta})$	SE
intercept	5.2958	$(.0898)^2 = .00806$	.3203	.000042	.00810	.0900
time	-.3346	$(.0670)^2 = .00449$	.2517	.000026	.00452	.0672
drug	.1086	$(.1029)^2 = .01059$	-.3987	.000065	.01066	.1032
drug $\times$ time	-.6868	$(.0776)^2 = .00602$	-.6348	.000165	.00619	.0786

$$\text{Augment} = \frac{n_d n_c}{N^3} (\hat{\beta}^D)^2, \text{ here, } \frac{n_d n_c}{N^3} = \frac{102 \times 335}{(437)^3} = .00040945$$

## Pattern-mixture averaged results - using drug-stratified proportions

### Placebo Intercept

$$\frac{70}{108}(5.22) + \frac{38}{108}(5.22 + 0.32) = 5.22 + \frac{38}{108}(0.32) = 5.33$$

*Completers*      *Dropouts*

### Placebo Time effect

$$\frac{70}{108}(-0.39) + \frac{38}{108}(-0.39 + 0.25) = -0.39 + \frac{38}{108}(0.25) = -0.30$$

*Completers*      *Dropouts*

### Drug Intercept difference

$$\frac{265}{329}(0.20) + \frac{64}{329}(0.20 - 0.40) = 0.20 + \frac{64}{329}(-0.40) = 0.12$$

*Completers*      *Dropouts*

### Drug Time difference

$$\frac{265}{329}(-0.54) + \frac{64}{329}(-0.54 - 0.64) = -0.54 + \frac{64}{329}(-0.64) = -0.66$$

*Completers*      *Dropouts*

$$\text{Calculation of } \hat{V}(\hat{\beta}) = \hat{V}(\hat{\beta})_F + \frac{n_d n_c}{N^3} (\hat{\beta}^D)^2$$

parameter	$\hat{\beta}$	$\hat{V}(\hat{\beta})_F$	$\hat{\beta}^D$	Augment	$\hat{V}(\hat{\beta})$	SE
intercept	5.3337	$(.0879)^2 = .00773$	.3203	.000217	.00795	.0891
time	-.3048	$(.0698)^2 = .00487$	.2517	.000134	.00500	.0707
drug	.1241	$(.1043)^2 = .01088$	-.3987	.000076	.01096	.1047
drug $\times$ time	-.6621	$(.0772)^2 = .00596$	-.6348	.000192	.00615	.0784

$$\text{Augment} = \frac{n_d n_c}{N^3} (\hat{\beta}^D)^2, \text{ where}$$

$$\frac{n_d n_c}{N^3} = \frac{38 \times 70}{(108)^3} = .00211159 \text{ for placebo}$$

$$\frac{n_d n_c}{N^3} = \frac{64 \times 265}{(329)^3} = .00047625 \text{ for drug}$$

NIMH Schizophrenia Study - IMPS79 across Time: MRM models

parameter	Ordinary			Pattern mixture			PM averaged		
	est	se	$p <$	est	se	$p <$	est	se	$p <$
Int $\beta_0$	5.348	.088	.0001	5.221	.108	.0001	5.334	.089	.0001
Drug $\beta_1$	.046	.101	.65	.202	.121	.096	.124	.105	.24
SWeek $\beta_2$	-.336	.068	.0001	-.393	.076	.0001	-.305	.071	.0001
Drug $\times$ SWeek $\beta_3$	-.641	.078	.0001	-.539	.086	.0001	-.662	.078	.0001
Int $\beta_0^D$				.320	.186	.086			
Drug $\beta_1^D$				-.399	.227	.079			
SWeek $\beta_2^D$				.252	.159	.115			
Drug $\times$ SWeek $\beta_3^D$				-.635	.196	.002			
Deviance	4649.0			4623.3					

## Mixed-effects selection models

These models have also been called

- random-coefficient selection models (Little, 95)
- random-effects-dependent models (Hogan & Laird, 97)
- shared parameter models (Wu & Carroll, 88; Ten Have *et al.*, 98)
  
- One specifies both a model for the longitudinal outcome and a model for the dropout (or missingness)
- Both models depend on random subject effects, most or all of which are shared by both models

**Longitudinal model** - ordinary MRM of  $\mathbf{y}_i$

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{v}_i + \boldsymbol{\varepsilon}_i$$

**Dropout model** - grouped/discrete time survival analysis of  $D_i$

$$\log(-\log(1 - P(D_i = j \mid D_i \geq j))) = \mathbf{W}_i\boldsymbol{\alpha} + \mathbf{v}_i\boldsymbol{\alpha}^*$$

$\mathbf{W}_i$  includes dropout predictors, some or all may be in  $\mathbf{X}_i$

To the extent that  $\boldsymbol{\alpha}^*$  are nonzero, this is a nonignorable model because missingness, here characterized simply as dropout time, is dependent on both  $\mathbf{y}_i^O$  and  $\mathbf{y}_i^M$  (via  $\mathbf{v}_i$ )

Treatment (denoted **Drug**) by last wave (denoted **Maxweek**)

Drug	Maxweek						Total
	1	2	3	4	5	6	
placebo	13 (.12)	5 (.05)	16 (.15)	2 (.02)	2 (.02)	70 (.65)	108
drug	24 (.07)	5 (.02)	26 (.08)	3 (.01)	6 (.02)	265 (.81)	329

⇒ dropout is more common among the placebo group

Pearson  $\chi^2$  test yields  $p < .025$ ;

Mantel-Haenszel  $\chi^2$  test for trend yields  $p < .0013$

## Mixed-effects selection model - Schiz study

Longitudinal model:

$$\text{IMPS79}_{ij} = \beta_0 + \beta_1 \text{Drug}_i + \beta_2 \text{SWeek}_j + \beta_3 (\text{Drug}_i \times \text{SWeek}_j) + v_{0i} + v_{1i} \text{SWeek}_j + \varepsilon_{ij}$$

or, after orthogonalizing the random effects

( $\mathbf{v}_i = \mathbf{S}\boldsymbol{\theta}_i$ , where  $\boldsymbol{\Sigma}_v = \mathbf{S}\mathbf{S}'$ , Cholesky factorization)

$$\begin{aligned} \text{IMPS79}_{ij} = & \beta_0 + \beta_1 \text{Drug}_i + \beta_2 \text{SWeek}_j + \beta_3 (\text{Drug}_i \times \text{SWeek}_j) \\ & + (\sigma_{v_0} + \sigma_{v_{01}}^* \text{SWeek}_j) \theta_{0i} + \sigma_{v_1}^* \text{SWeek}_j \theta_{1i} + \varepsilon_{ij} \end{aligned}$$

with  $\sigma_{v_{01}}^* = \sigma_{v_{01}} / \sigma_{v_0}$  and  $\sigma_{v_1}^* = \sqrt{\sigma_{v_1}^2 - \sigma_{v_{01}}^2} / \sigma_{v_0}$

Dropout model:

$$\begin{aligned} \log(-\log(1 - P(D_i = j \mid D_i \geq j))) &= \alpha_{0j} + \alpha_1 \mathbf{Drug}_i + \alpha_2 \theta_{0i} + \alpha_3 \theta_{1i} \\ &\quad + \alpha_4 (\mathbf{Drug}_i \times \theta_{0i}) + \alpha_5 (\mathbf{Drug}_i \times \theta_{1i}) \end{aligned}$$

or as

$$\begin{aligned} P(D_i \leq j) &= 1 - \exp(-\exp(\alpha_{0j} + \alpha_1 \mathbf{Drug}_i + \alpha_2 \theta_{0i} + \alpha_3 \theta_{1i} \\ &\quad + \alpha_4 (\mathbf{Drug}_i \times \theta_{0i}) + \alpha_5 (\mathbf{Drug}_i \times \theta_{1i}))) \end{aligned}$$

- random effects are summaries of a person's observed *and unobserved*  $\mathbf{y}$  data
- this shared parameter model is a nonignorable model if  $\alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$  is rejected
- test of whether *a particular model of ignorability is reasonable vs a particular model of nonignorability (i.e., not a general test of ignorability)*

## Separate and shared parameter models

parameter	Separate (deviance = 5380.2)			Shared (deviance = 5350.1)		
	ML est	std error	<i>p</i> -value	ML est	std error	<i>p</i> -value
<u>Outcome</u>						
intercept $\beta_0$	5.348	.088	.0001	5.320	.088	.0001
Drug $\beta_1$	.046	.101	.65	.088	.102	.87
SWeek $\beta_2$	-.336	.068	.0001	-.272	.073	.0002
Drug $\times$ Sweek $\beta_3$	-.641	.078	.0001	-.737	.083	.0001
<u>Dropout</u>						
Drug $\alpha_1$	-.693	.205	.0008	-.703	.301	.02
Random intercept $\alpha_2$				.447	.333	.18
Random slope $\alpha_3$				.891	.467	.06
Drug $\times$ intercept $\alpha_4$				-.592	.398	.14
Drug $\times$ slope $\alpha_5$				-1.638	.536	.003

- separate parameter model yields identical results as running these two models, one for  $\mathbf{y}_i$  and one for  $D_i$ , separately
- shared parameter model fits better,  $\chi_4^2 = 30.1, p < .0001$
- for longitudinal component, conclusions are same as MAR model
- marginally significant slope: for the placebo group there is a tendency to dropout as the slope increases
- significant negative **Drug**  $\times$  slope: the slope effect is opposite for the drug group; drug patients with more negative slopes (*i.e.*, greater improvement) are more likely to drop out

NIMH Schizophrenia Study: Severity across Time  
 ML Estimates (se) *random intercept and slope models*

	<i>Completers</i> <i>N = 335</i>	<i>All cases</i> <i>N = 437</i>	<i>Shared</i> <i>Parameter</i> <i>N = 437</i>	<i>Pattern</i> <i>Mixture</i> <i>N = 437</i>
intercept	5.221 (.109)	5.348 (.088)	5.320 (.088)	5.334 (.089)
Drug (0=P; 1=D)	0.202 (.123)	0.046 (.101)	0.088 (.102)	0.124 (.105)
Time (sqrt wk)	-0.393 (.073)	-0.336 (.068)	-0.272 (.073)	-0.305 (.071)
Drug by Time	-0.539 (.083)	-0.641 (.078)	-0.737 (.083)	-0.662 (.078)

## Conclusions

- Mixed-effects regression models (MRMs) useful for incomplete longitudinal data
  - can handle subjects measured incompletely or at different timepoints
  - missing data assumed MAR
    - \* dependent on covariates *and*
    - \* available data on dependent variable

- Mixed-effects pattern-mixture and selection (*i.e.*, shared parameter) models augment MRM

### Pattern-mixture

- adds missing-data pattern as between-subjects factor
- assesses degree to which “missingness” influences outcomes
- assesses degree to which “missingness” interacts with model terms (*i.e.*, intervention group, intervention group by time)

### Selection

- missingness in terms of important covariates
- missingness in terms of (shared) random subject effects

⇒ Does not invent data, attempts to maximize information obtained from *available* data

## SAS MIXED code - SCHIZPM5b.SAS

---

```
DATA one; INFILE 'c:\mixdemo\schizrep.dat';
INPUT id imps79 week drug sex ;

/* the coding for the variables is as follows:
id = subject id number
imps79 = overall severity (1=normal, ..., 7=among the most extremely ill)
week = 0,1,2,3,4,5,6 (most of the obs. are at weeks 0,1,3, and 6)
drug 0=placebo 1=drug (chlorpromazine, fluphenazine, or thioridazine)
sex 0=female 1=male */

/* compute the square root of week to linearize relationship */
sweek = SQRT(week);

/* calculate the maximum value of week for each subject
(suppress the printing of the output for this procedure) */
PROC MEANS NOPRINT; CLASS id; VAR week; OUTPUT OUT=two MAX=maxweek;
RUN;

/* determine if a subject has data at week 6
dropout = 0 (for completers) or = 1 (for dropouts) */
DATA three; SET two;
dropout=0;
IF maxweek LT 6 THEN dropout=1;
```

```

/* dataset with all subjects (adding the dropout variable) */
DATA four; MERGE one three; BY id; IF id NE .;

/* random intercept and trend model */
PROC MIXED DATA=four METHOD=ML COVTEST;
CLASS id;
MODELimps79 = sweek drug sweek*drug / SOLUTION;
RANDOM INTERCEPT sweek /SUB=id TYPE=UN G GCORR;
RUN;

/* pattern-mixture random intercept and trend model
/* using marginal dropout proportion to estimate averaged results */
PROC MIXED DATA=four METHOD=ML COVTEST;
CLASS id;
MODELimps79 = sweek drug sweek*drug dropout dropout*sweek
              dropout*drug dropout*drug*sweek / SOLUTION;
RANDOM INTERCEPT sweek /SUB=id TYPE=UN G GCORR;
ESTIMATE 'avg int' INTERCEPT 1 sweek 0 drug 0 sweek*drug 0 dropout .2334
          dropout*sweek 0 dropout*drug 0 dropout*drug*sweek 0;
ESTIMATE 'avg sweek' INTERCEPT 0 sweek 1 drug 0 sweek*drug 0 dropout 0
          dropout*sweek .2334 dropout*drug 0 dropout*drug*sweek 0;
ESTIMATE 'avg drug' INTERCEPT 0 sweek 0 drug 1 sweek*drug 0 dropout 0
          dropout*sweek 0 dropout*drug .2334 dropout*drug*sweek 0;
ESTIMATE 'avg sweek*drug' INTERCEPT 0 sweek 0 drug 0 sweek*drug 1 dropout 0
          dropout*sweek 0 dropout*drug 0 dropout*drug*sweek .2334;
RUN;

```

```
/* pattern-mixture random intercept and trend model
/* using drug-specific dropout proportions to estimate averaged results */

PROC MIXED DATA=four METHOD=ML COVTEST;
CLASS id;
MODEL imps79 = sweek drug sweek*drug dropout dropout*sweek
              dropout*drug dropout*drug*sweek / SOLUTION;
RANDOM INTERCEPT sweek /SUB=id TYPE=UN G GCORR;
ESTIMATE 'avg int' INTERCEPT 1 sweek 0 drug 0 sweek*drug 0 dropout .35185
          dropout*sweek 0 dropout*drug 0 dropout*drug*sweek 0;
ESTIMATE 'avg sweek' INTERCEPT 0 sweek 1 drug 0 sweek*drug 0 dropout 0
          dropout*sweek .35185 dropout*drug 0 dropout*drug*sweek 0;
ESTIMATE 'avg drug' INTERCEPT 0 sweek 0 drug 1 sweek*drug 0 dropout 0
          dropout*sweek 0 dropout*drug .19453 dropout*drug*sweek 0;
ESTIMATE 'avg sweek*drug' INTERCEPT 0 sweek 0 drug 0 sweek*drug 1 dropout 0
          dropout*sweek 0 dropout*drug 0 dropout*drug*sweek .19453;

RUN;
```

---

## SAS code - SCHZ2MODS.SAS

---

```
TITLE1 'shared parameter model of time to dropout and imps79 across time';
DATA one; INFILE 'c:\mixdemo\schizrep.dat'; INPUT id imps79 week drug sex ;

/* The coding for the variables is as follows:
id = subject id number
imps79 = overall severity (1=normal, ..., 7=most extremely ill)
week = 0,1,2,3,4,5,6 (most of the obs. are at weeks 0,1,3, and 6)
drug 0=placebo 1=drug (chlorpromazine, fluphenazine, or thioridazine)
sex 0=female 1=male */

/* compute the square root of week to linearize relationship */
sweek = SQRT(week);

/* calculate the maximum value of week for each subject
and get drug in this aggregated dataset too */
PROC MEANS NOPRINT; CLASS id; VAR week drug;
OUTPUT OUT=two MAX(week drug)=maxweek drug;

/* setting up imps79 across time and maxweek as one outcome vector */
DATA daty; SET one; outcome = imps79; ind = 0;
DATA datr; SET two; outcome = maxweek; ind = 1; IF id NE .;
DATA all; SET daty datr; BY id;
```

---

- uppercase represents specific SAS syntax; lowercase represents user-defined
- the SAS dataset `all` includes the  $(n_i + 1) \times 1$  outcome vector  $\mathbf{y}_i^*$ , named `outcome`, which contains  $\mathbf{y}_i$  as its first  $n_i$  elements and  $D_i$  as its final element
- `ind` with values of 0 or 1, is also defined; this variable will be used to distinguish between the  $\mathbf{y}_i$  and  $D_i$  elements

---

```
/* random trend model for imps79 */
PROC MIXED DATA=daty METHOD=ML COVTEST;
CLASS id;
MODEL outcome = drug sweek sweek*drug / SOLUTION;
RANDOM INTERCEPT sweek /SUB=id TYPE=UN G GCORR;

/* model for time to dropout - grouped-time proportional hazards model */
PROC LOGISTIC DATA=datr;
MODEL outcome = drug / LINK = CLOGLOG;
```

---

---

```

/* separate modeling of imps79 and time to dropout */
PROC NL MIXED DATA=all;
PARMS b0=5.35 b1=.05 b2=-.34 b3=-.64 sde=.76 v0=.61 v01=.02 v1=.49
a1=-.69 i1=-1.95 i2=-1.70 i3=-.99 i4=-.84 i5=-.69;
IF (ind = 0) THEN
DO;
  z = (outcome - (b0 + b1*drug + b2*sweek + b3*drug*sweek
    + (v0 + v01*sweek/v0)*u1
    + SQRT(v1*v1 - (v01*v01)/(v0*v0))*sweek*u2));
  p = (1 / SQRT(2*3.14159*sde*sde)) * EXP(-.5 * (z*z) / (sde*sde));
END;
IF (ind = 1) THEN
DO;
  z = a1*drug;
  IF (outcome=1) THEN
    p = 1 - EXP(0 - EXP(i1+z));
  ELSE IF (outcome=2) THEN
    p = (1 - EXP(0 - EXP(i2+z))) - (1 - EXP(0 - EXP(i1+z)));
  ELSE IF (outcome=3) THEN
    p = (1 - EXP(0 - EXP(i3+z))) - (1 - EXP(0 - EXP(i2+z)));
  ELSE IF (outcome=4) THEN
    p = (1 - EXP(0 - EXP(i4+z))) - (1 - EXP(0 - EXP(i3+z)));
  ELSE IF (outcome=5) THEN
    p = (1 - EXP(0 - EXP(i5+z))) - (1 - EXP(0 - EXP(i4+z)));
  ELSE IF (outcome=6) THEN
    p = 1 - (1 - EXP(0 - EXP(i5+z)));
END;
IF (p > 1e-8) THEN ll = LOG(p); else ll = -1e100;
MODEL outcome ~ GENERAL(ll);
RANDOM u1 u2 ~ NORMAL([0,0], [1,0,1]) SUBJECT=id;

```

---

---

```

/* shared parameter model of imps79 and time to dropout */
PROC NL MIXED DATA=all;
PARMS b0=5.35 b1=.05 b2=-.34 b3=-.64 sde=.76 v0=.61 v01=.02 v1=.49
a1=-.69 a2=0 a3=0 a4=0 a5=0 i1=-1.95 i2=-1.70 i3=-.99 i4=-.84 i5=-.69;
IF (ind = 0) THEN
DO;
  z = (outcome - (b0 + b1*drug + b2*sweek + b3*drug*sweek
    + (v0 + v01*sweek/v0)*u1
    + SQRT(v1*v1 - (v01*v01)/(v0*v0))*sweek*u2));
  p = (1 / SQRT(2*3.14159*sde*sde)) * EXP(-.5 * (z*z) / (sde*sde));
END;
IF (ind = 1) THEN
DO;
  z = a1*drug + a2*u1 + a3*u2 + a4*u1*drug + a5*u2*drug;
  IF (outcome=1) THEN
    p = 1 - EXP(0 - EXP(i1+z));
  ELSE IF (outcome=2) THEN
    p = (1 - EXP(0 - EXP(i2+z))) - (1 - EXP(0 - EXP(i1+z)));
  ELSE IF (outcome=3) THEN
    p = (1 - EXP(0 - EXP(i3+z))) - (1 - EXP(0 - EXP(i2+z)));
  ELSE IF (outcome=4) THEN
    p = (1 - EXP(0 - EXP(i4+z))) - (1 - EXP(0 - EXP(i3+z)));
  ELSE IF (outcome=5) THEN
    p = (1 - EXP(0 - EXP(i5+z))) - (1 - EXP(0 - EXP(i4+z)));
  ELSE IF (outcome=6) THEN
    p = 1 - (1 - EXP(0 - EXP(i5+z)));
END;
IF (p > 1e-8) THEN ll = LOG(p); else ll = -1e100;
MODEL outcome ~ GENERAL(ll);
RANDOM u1 u2 ~ NORMAL([0,0], [1,0,1]) SUBJECT=id;

```

---