Contemporary Issues in Clinical Trials Methods
Longitudinal Data Analysis Part II

Michael Donohue

Alzheimer's Therapeutic Research Institute
Department of Neurology
University of Southern California

July 14, 2017

Key features of longitudinal models

- Last session we discussed a linear mixed effect model (LME) which treats time as a continuous variable
- This is one type of longitudinal model among many types
- Three key features of longitudinal models:
  - **Mean structure**: Governs shape of mean trajectory
  - **Variance-covariance structure**: Governs within-subject correlation
  - **Baseline assessment**: Treated as covariate or outcome

Taxonomy of longitudinal models

	MMRM	**LDA**	**cLDA**
Mean structure	cat. time	cat. or cts. time	cat. or cts. time
Variance-covariance	correlated residuals	correlated residuals or random effects	correlated residuals or random effects
Baseline assessment	as covariate	as outcome; different means per group	as outcome; common mean per group

Abbreviations: MMRM, Mixed Model of Repeated Measures; LDA, Longitudinal Data Analysis; cLDA, Constrained LDA; cat., categorical; cts., continuous.

Linear mixed-effects model (LME)

$$\text{ADAS}_{ij} = X_i\beta + b_{0i} + t_{ij}b_{1i} + \varepsilon_{ij}$$

- X_i: covariates for subject $i = 1, \ldots, 400$
- β: population level "fixed effects"
- $b_i \sim \mathcal{N}(0, D)$: subject-specific "random effects" for subject $i = 1, \ldots, 400$
- $(\varepsilon_{i1}, \ldots, \varepsilon_{i4}) \sim \mathcal{N}(0, \Sigma)$: vector of "residuals" for subject $i = 1, \ldots, 400$
- D, Σ: "variance components"

$b_1, \ldots, b_N, \varepsilon_1, \ldots, \varepsilon_N$ are assumed independent

Linear mixed-effects model (LME)

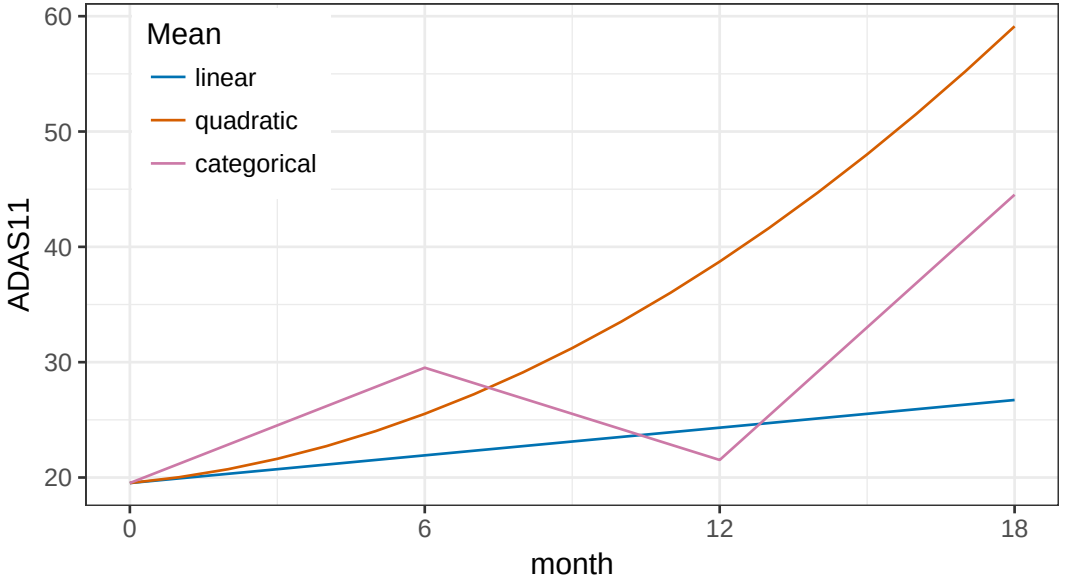
$$\text{ADAS}_{ij} = X_i\beta + b_{0i} + t_{ij}b_{1i} + \varepsilon_{ij}$$

- "Mean structure" (fixed effects): $X_i\beta$
  - linear, quadratic, or categorical time; and
  - other fixed-effect covariates (age, education, gender, treatment)
- "Variance structure" (random effects & residuals): $b_{0i} + t_{ij}b_{1i} + \varepsilon_{ij}$
  - random intercept, random intercept & slope; or
  - impose variance-covariance (variance and correlation) structure on residuals $(\varepsilon_{i1}, \ldots, \varepsilon_{i4}) \sim \mathcal{N}(0, \Sigma)$

"Mixed Model of Repeated Measures" (MMRM)

- A popular marginal model, "Mixed Model of Repeated Measures" (MMRM), makes no (or very general) assumptions about the mean and variance/covariance structure.
- Unstructured mean: time is treated as categorical, so no linear (or quadratic, etc.) trend is assumed
- Unstructured variance/covariance: includes parameters for variance at each visit ("heterogeneous") and visit-to-visit correlation
- A repeated measures extension of ANCOVA (instead of one post-baseline assessment, there are several)
- MMRM is not actually a mixed-effects model (no random effects)!
- Usually treats change from baseline as outcome variable, and baseline as covariate (like ANCOVA).

Mean structure examples



Variance-covariance structure

$$\text{ADAS}_{ij} = X_i\beta + b_{0i} + t_{ij}b_{1i} + \varepsilon_{ij}$$

- *Vanilla* LME assumes *homoscedastic* (homogenous/constant variance), independent residuals
- MMRM drops the b terms and can assume *heteroscedastic* (heterogeneous variance), correlated residuals
  - $(\varepsilon_{i1}, \ldots, \varepsilon_{i4}) \sim \mathcal{N}(0, \Sigma)$
  - $\Sigma = \sigma^2 VCV$, where
  - σ^2 is variance scale parameter
  - V is diagonal matrix of standard deviation weights
  - C is correlation matrix

Variance-covariance structure

$$(\varepsilon_{i1}, \ldots, \varepsilon_{i4}) \sim \mathcal{N}(0, \Sigma), \quad \Sigma = \sigma^2 VCV$$

Examples:

Heterogeneous variance

$$V = \begin{array}{c|cccc} & 0 & 6 & 12 & 18 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 6 & 0 & 2 & 0 & 0 \\ 12 & 0 & 0 & 7 & 0 \\ 18 & 0 & 0 & 0 & 10 \end{array}$$

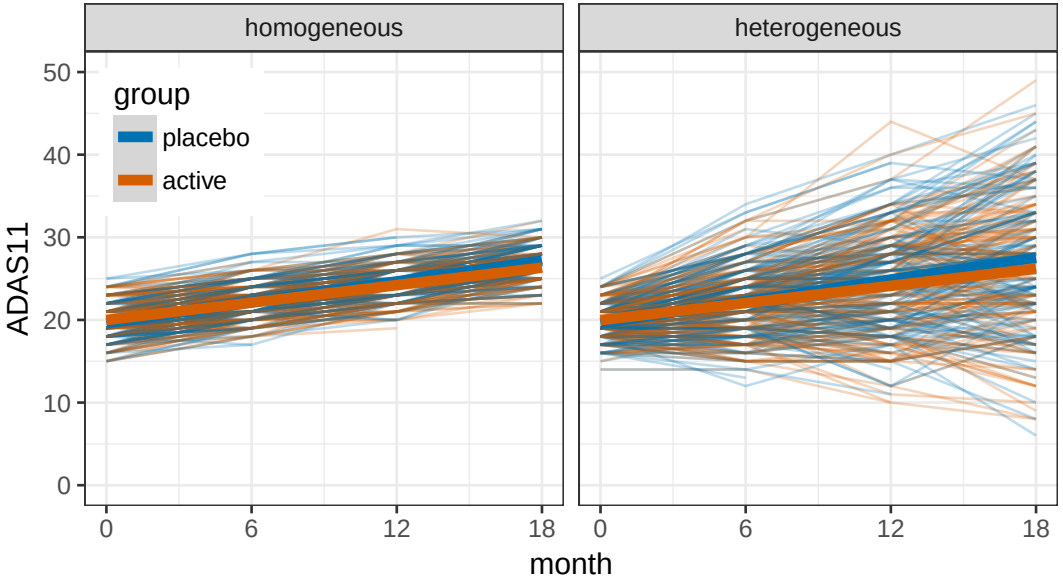
Exchangeable/Compound Symmetric

$$C = \begin{array}{c|cccc} & 0 & 6 & 12 & 18 \\ \hline 0 & 1.0 & 0.2 & 0.2 & 0.2 \\ 6 & 0.2 & 1.0 & 0.2 & 0.2 \\ 12 & 0.2 & 0.2 & 1.0 & 0.2 \\ 18 & 0.2 & 0.2 & 0.2 & 1.0 \end{array}$$

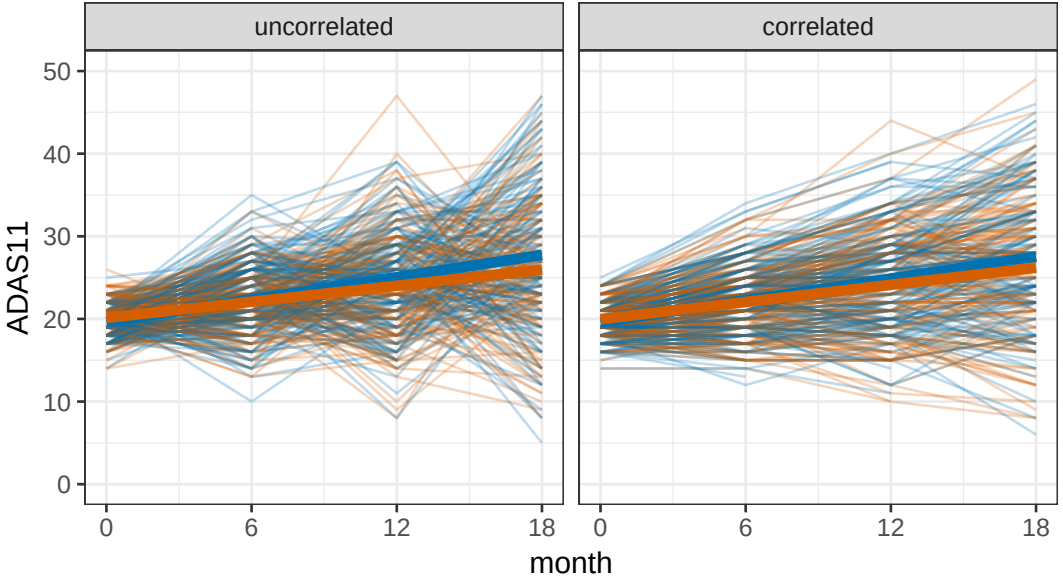
Unstructured/Symmetric

$$C = \begin{array}{c|cccc} & 0 & 6 & 12 & 18 \\ \hline 0 & 1.0 & 0.3 & 0.2 & 0.1 \\ 6 & 0.3 & 1.0 & 0.4 & 0.2 \\ 12 & 0.2 & 0.4 & 1.0 & 0.2 \\ 12 & 0.1 & 0.2 & 0.2 & 1.0 \end{array}$$

Homogeneous vs heterogeneous variance structure



Uncorrelated vs correlated residuals



Let's simulate categorical time data from a hypothetical clinical trial...

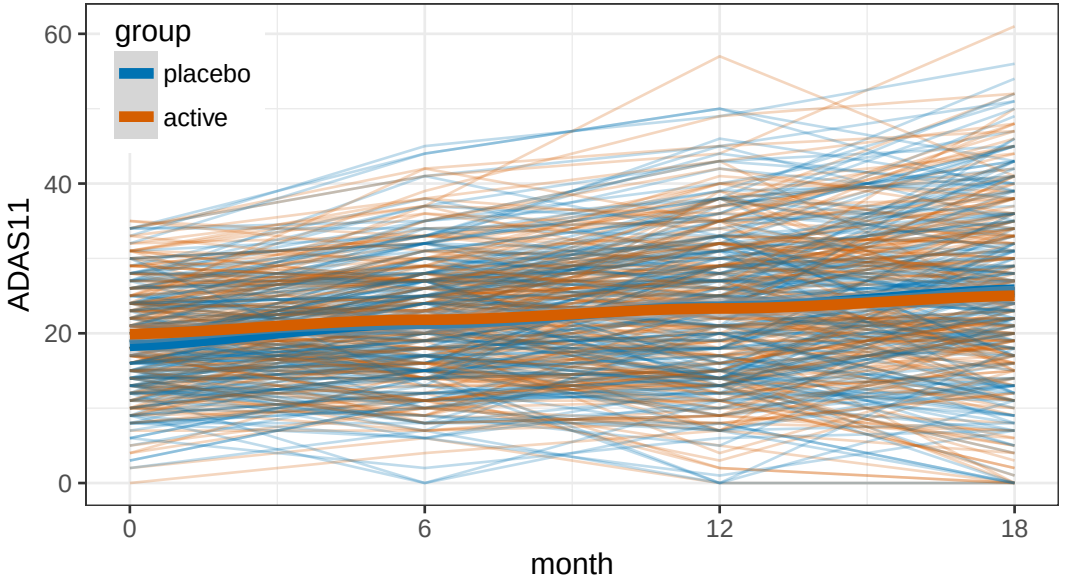
These are all estimates required:

```
Beta <- c(
   '(Intercept)'= 19.8, # mean ADAS at baseline
        'female'=-0.51, # female perform better
         'age_c'= 0.04, # worse change for older at baseline (age mean centered)
            'm6'= 2.23, # worsening at month 6 in pbo
           'm12'= 4.46, # worsening at month 12 in pbo
           'm18'= 7.31, # worsening at month 18 in pbo
     'm6:active'=-0.20, # relative improvement at month 6 with treatment
    'm12:active'=-0.70, # relative improvement at month 12 with treatment
    'm18:active'=-1.75) # relative improvement at month 18 with treatment

# other design parameters
months <- c(0, 6, 12, 18)
n <- 200 # per group
attrition_rate <- 0.40/18 # approx per month

# var-cov parameters
SD <- 6.77                          # standard deviation scale parameter
vv <- diag(c(1, 1.2, 1.5, 1.8))          # hetergeneous variance weight matrix
cc <- matrix(0.75, nrow=4, ncol=4)   # correlation matrix
```

The pseudo categorical time data

	id	ADAS11.m0	active	female	age	censor	age_c	month	residual	visNo	m0	m6	m12	m18	ADAS11
1	1	20	1	0	66.8	101.2	-8.49	6	3.86	1	0	1	0	0	25
2	1	20	1	0	66.8	101.2	-8.49	12	5.63	2	0	0	1	0	29
3	1	20	1	0	66.8	101.2	-8.49	18	1.69	3	0	0	0	1	27
4	2	15	0	0	63.4	22.5	-11.90	6	3.11	1	0	1	0	0	25
5	2	15	0	0	63.4	22.5	-11.90	12	-1.75	2	0	0	1	0	22
6	2	15	0	0	63.4	22.5	-11.90	18	-6.27	3	0	0	0	1	20

The pseudo categorical time data



Fitting MMRM in R

```
# Symmetric correlation, hetergeneous variance
MMRMsymHet <- gls(ADAS11.ch ~
  -1+ADAS11.m0+female+age_c+(m6+m12+m18)+(m6+m12+m18):active,
  data=trial_mmrm, correlation = corSymm(form = ~ visNo | id),
  weights = varIdent(form = ~ 1 | m) )

# Compound Symmetric correlation, hetergeneous variance
MMRMcompSymHet <- gls(ADAS11.ch ~
  -1+ADAS11.m0+female+age_c+(m6+m12+m18)+(m6+m12+m18):active,
  data=trial_mmrm, correlation = corCompSymm(form = ~ visNo | id),
  weights = varIdent(form = ~ 1 | m) )

# see ?corClasses and ?varClasses for more options
```

MMRM summaries in R

```
summary(MMRMsymHet)
```

```
Correlation Structure: General
 Formula: ~visNo | id
 Parameter estimate(s):
 Correlation:
  1     2
2 0.457
3 0.269 0.406
Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | m
 Parameter estimates:
   6   12   18
1.00 1.29 1.60
```

MMRM summaries in R

```
Coefficients:
             Value Std.Error t-value p-value
ADAS11.m0   0.0187    0.0396    0.47  0.6372
female      0.2335    0.5375    0.43  0.6641
age_c       0.0112    0.0357    0.31  0.7532
m6          2.7454    0.8762    3.13  0.0018 **
m12         4.7068    0.9520    4.94  <0.001 ***
m18         8.0360    1.0791    7.45  <0.001 ***
m6:active  -1.3766    0.5775   -2.38  0.0173 *
m12:active -1.9003    0.7841   -2.42  0.0156 *
m18:active -3.5408    1.0404   -3.40  <0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 5.4
```

Fitting cLDA in R

```
# Symmetric correlation, hetergeneous variance
cLDAsymHet <- gls(ADAS11 ~
  -1+female+age_c+m0+(m6+m12+m18)+(m6+m12+m18):active,
  data=trial_obs, correlation = corSymm(form = ~ visNo | id),
  weights = varIdent(form = ~ 1 | m) )

# Compound Symmetric correlation, hetergeneous variance
cLDAcompSymHet <- gls(ADAS11 ~
  -1+female+age_c+m0+(m6+m12+m18)+(m6+m12+m18):active,
  data=trial_obs, correlation = corCompSymm(form = ~ visNo | id),
  weights = varIdent(form = ~ 1 | m) )

# see ?corClasses and ?varClasses for more options
```

cLDA summaries in R

```
Correlation Structure: General
 Formula: ~visNo | id
 Parameter estimate(s):
 Correlation:
  1     2     3
2 0.763
3 0.751 0.779
4 0.751 0.718 0.738
Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | m
 Parameter estimates:
   0    6   12   18
1.00 1.21 1.53 1.86
```

MMRM summaries in R

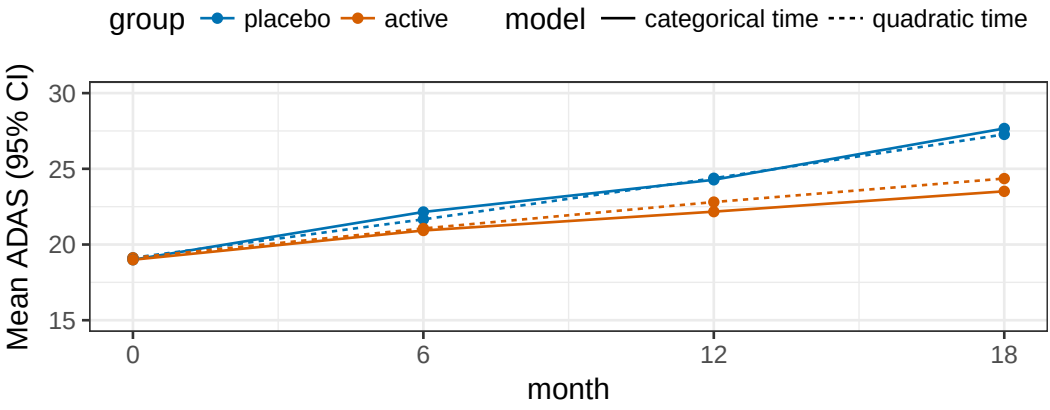
```
Coefficients:
             Value Std.Error t-value p-value
female     -0.2519    0.6527   -0.39  0.6996
age_c       0.0600    0.0424    1.42  0.1569
m0         19.2454    0.4942   38.94  <0.001 ***
m6         22.4019    0.6212   36.06  <0.001 ***
m12        24.5268    0.7598   32.28  <0.001 ***
m18        27.9168    0.9279   30.09  <0.001 ***
m6:active  -1.2232    0.5679   -2.15  0.0314 *
m12:active -2.1045    0.7744   -2.72  0.0067 **
m18:active -4.1514    0.9996   -4.15  <0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 6.83
```

Fitting continuous time cLDA in R

```
# Linear time, random intercept
cLDAlin <- lme(ADAS11 ~
  female + age_c + month + month:active,
  data=trial_obs, random = ~1|id)

# Quadratic time, random intercept
cLDAquad <- lme(ADAS11 ~
  female + age_c + (month + I(month^2)) + (month + I(month^2)):active,
  data=trial_obs, random = ~1|id)
```

Modeled mean profiles



active	200	174	152	138
placebo	200	180	158	132

Modeled group contrasts for quadratic and categorical time models

```
 Simultaneous Tests for General Linear Hypotheses

Fit: lme.formula(fixed = ADAS11 ~ female + age_c + (month + I(month^2)) +
    (month + I(month^2)):active, data = trial_obs, random = ~1 |
    id)

Linear Hypotheses:
         Estimate Std. Error z value Pr(>|z|)
m6 == 0    -0.602      0.572   -1.05    0.506
m12 == 0   -1.574      0.683   -2.31    0.047 *
m18 == 0   -2.917      0.774   -3.77   <0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

Modeled group contrasts for quadratic and categorical time models

```
 Simultaneous Tests for General Linear Hypotheses

Fit: gls(model = ADAS11 ~ -1 + female + age_c + m0 + (m6 + m12 + m18) +
    (m6 + m12 + m18):active, data = trial_obs, correlation = corSymm(form =
    id), weights = varIdent(form = ~1 | m))

Linear Hypotheses:
         Estimate Std. Error z value Pr(>|z|)
m6 == 0    -1.223      0.568   -2.15    0.085 .
m12 == 0   -2.105      0.774   -2.72    0.019 *
m18 == 0   -4.151      1.000   -4.15   <0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

Taxonomy of longitudinal models

	MMRM	**LDA**	**cLDA**
Mean structure	cat. time	cat. or cts. time	cat. or cts. time
Variance-covariance	correlated residuals	correlated residuals or random effects	correlated residuals or random effects
Baseline assessment	as covariate	as outcome; different means per group	as outcome; common mean per group

Abbreviations: MMRM, Mixed Model of Repeated Measures; LDA, Longitudinal Data Analysis; cLDA, Constrained LDA; cat., categorical; cts., continuous.

MMRM vs cLDA

- Both (can use) categorical time
- MMRM has *change from baseline* as outcome, baseline assessment as covariate
  - → need follow-up data to calculate change
  - → modified Intention-to-Treat (mITT) population (i.e. randomized and at least one follow-up assessment)
- cLDA uses raw assessment scores as outcome; baseline assessment as outcome, constrains two groups to have same mean
  - → only need one observation to enter into analysis
  - → "*unmodified*" Intention-to-Treat population (i.e. randomized with at least one assessment)
- In our simulated trial it is a difference of $n = 46$ subjects

Abbreviations: MMRM, Mixed Model of Repeated Measures; LDA, Longitudinal Data Analysis; cLDA, Constrained LDA.

MMRM vs cLDA

From Lu (2010):

- cLDA more powerful than MMRM (aka "longitudinal ANCOVA") when baseline assessments are missing
- cLDA advantage is greater when correlation between baseline and follow-up is weaker

Caveat: Implicitly assumes data Missing at Random.

Lu, K. (2010). On efficiency of constrained longitudinal data analysis versus longitudinal analysis of covariance. *Biometrics*, 66(3), 891-896.

Missing Data: Notation

- Measurement: Y_{ij} (e.g. ADAS_{ij})
- Covariates: X_i (e.g. treatment, gender, age, ...)
- Missingness indicator:

$$R_{ij} = \begin{cases} 1 \text{ if } Y_{ij} \text{ is observed} \\ 0 \text{ otherwise} \end{cases}$$

- Let $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{in_i})' = (\mathbf{Y}_i^o, \mathbf{Y}_i^m)$, where

$$\begin{cases} \mathbf{Y}_i^o \text{ observed } Y_{ij} \\ \mathbf{Y}_i^m \text{ un-observed } Y_{ij} \end{cases}$$

- D_i is time of dropout
- θ parameters that control $\mathbf{Y}_i$ (e.g. effects for treatment, gender, age, ...)
- ψ parameters that control $\mathbf{R}_i$ (e.g. treatment effect, $\mathbf{Y}_i^m$, ...)

The notation is supposed to be a helpful shorthand. If it's not helpful, don't worry about it!

Taxonomy of missing data mechanism

Missing data mechanisms are defined in terms of the assumed distribution function, F, or "data generating mechanism," for the missingness indicator R_i

Term	Notation	Missingness depends on:
Missing **Completely** at Random	$\mathbf{R}_i \sim F(X_i, \psi)$	covariates
Missing at Random	$\mathbf{R}_i \sim F(X_i, \mathbf{Y}_i^o, \psi)$	covariates observed assessments
Missing **Not** at Random	$\mathbf{R}_i \sim F(X_i, \mathbf{Y}_i^o, \mathbf{Y}_i^m, \psi)$	covariates observed assessments **unobserved** assessments

We'll unpack these a bit ...

Missing Completely at Random (MCAR)

$\mathbf{R}_i \sim F(X_i, \psi)$: Missingness (may) depend only on observed covariates (X_i)
Appropriate methods:

- Complete Case (e.g. ANCOVA) (?)
- Last Observation Carried Forward (LOCF) (?)
- Single imputation (?)

Missing at Random (MAR)

$\mathbf{R}_i \sim F(X_i, \mathbf{Y}_i^o, \psi)$: Missingness (may) depend on observed covariates (X_i) and/or observed outcomes ($\mathbf{Y}_i^o$)
Appropriate methods:

- **Direct likelihood (e.g. mixed-effects models, MMRM, cLDA)!**,
- Multiple imputation
- Weighted generalized estimating equations (WGEE)

Missing Not at Random (MNAR)

$\mathbf{R}_i \sim F(X_i, \mathbf{Y}_i^o, \mathbf{Y}_i^m, \psi)$: Missingness (may) depend on observed covariates (X_i), observed outcomes ($\mathbf{Y}_i^o$), and unobserved outcomes ($\mathbf{Y}_i^m$)
Appropriate sensitivity analyses (?):

- selection models: $f(\mathbf{Y}_i|X_i, \theta)f(\mathbf{R}_i|X_i, \mathbf{Y}_i^o, \mathbf{Y}_i^m, \psi)$
- pattern-mixture models (e.g. "tipping point" stress test): $f(\mathbf{Y}_i|X_i, \mathbf{R}_i, \theta)f(\mathbf{R}_i|X_i, \psi)$
- shared-parameter models: $f(\mathbf{Y}_i|X_i, \mathbf{b}_i, \theta)f(\mathbf{R}_i|X_i, \mathbf{b}_i, \psi)$

All of these approaches make untestable assumptions about missingness.
We can never completely rule out MNAR, since, if it exists, it depends on variables that we do not observe.

Quiz

MCAR, MAR, or MNAR?

1. Missingness at month 18 only depends on treatment assignment.
2. Missingness at month 18 only depends on treatment assignment and ADAS at month 12.
3. Missingness at month 18 only depends on treatment assignment and ADAS at month 18.

Missing data: bottom line

- Missing Not at Random is impossible to rule out. The best we can do is *sensitivity analyses* or apply models that make strong untestable assumptions about missingness mechanism.
- Missing Completely At Random is an unrealistic and unnecessary assumption.
- Missing at Random is a more reasonable assumption, and we have reliable methods that are robust in this case.

Direct likelihood (mixed-effects) is recommended. Complete case (ANCOVA), LOCF, and single imputation should be avoided.

Multiple imputation (MI)

Basic steps:

1. Create multiple complete versions of the data with imputed plausible values
2. Analyze each complete version with standard methods (e.g. ANCOVA)
3. Combine the results

Comments:

- MI requires many more modeling decisions than direct likelihood methods (e.g. number of imputations, imputation methods, ...)
- CAUTION: Lots of nuance and complexity
- Usually reserved for *sensitivity analyses*

1. Create multiple complete versions

```
trial_imp <- mice(trial_wide, predictorMatrix=predictorMatrix, seed = 2017
```

```
head(trial_wide)    # raw data with missing values:
```

	id	female	age	age_c	active	group	ADAS11.m0	ADAS11.m6	ADAS11.m12	ADAS11.m18
1	1	0	66.8	-8.489	1	active	20	25	29	27
2	2	0	63.4	-11.904	0	placebo	15	25	22	20
3	3	0	80.6	5.291	0	placebo	16	23	16	27
4	4	1	75.4	0.159	1	active	22	24	NA	NA
5	5	0	70.6	-4.647	1	active	29	33	38	41
6	6	1	76.4	1.148	1	active	16	20	22	22

```
head(complete(trial_imp)) # first complete version:
```

	id	female	age	age_c	active	group	ADAS11.m0	ADAS11.m6	ADAS11.m12	ADAS11.m18
1	1	0	66.8	-8.489	1	active	20	25	29	27
2	2	0	63.4	-11.904	0	placebo	15	25	22	20
3	3	0	80.6	5.291	0	placebo	16	23	16	27
4	4	1	75.4	0.159	1	active	22	24	29	29
5	5	0	70.6	-4.647	1	active	29	33	38	41
6	6	1	76.4	1.148	1	active	16	20	22	22

2. Analyze each complete version

```
 ## summary of imputation 1 :

Call:
lm(formula = ADAS11.m18 ~ active * center(ADAS11.m0))

Residuals:
    Min      1Q  Median      3Q     Max
-18.257  -5.851  -0.822   5.168  30.277

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)              27.5556     0.5970   46.15  < 2e-16 ***
active                   -3.9134     0.8442   -4.64  4.8e-06 ***
center(ADAS11.m0)         1.4284     0.0883   16.17  < 2e-16 ***
active:center(ADAS11.m0) -0.0099     0.1236   -0.08     0.94
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.39 on 396 degrees of freedom
Multiple R-squared:  0.574,Adjusted R-squared:  0.571
F-statistic:  178 on 3 and 396 DF,  p-value: <2e-16


 ## summary of imputation 2 :
```

3. Combine the results

```
                              est      se       t    df Pr(>|t|)
(Intercept)               27.7115  0.6520 42.5041  99.3   <0.001 ***
active                    -3.9801  0.9171 -4.3397 107.2   <0.001 ***
center(ADAS11.m0)          1.4180  0.0945 15.0036 136.2   <0.001 ***
active:center(ADAS11.m0) -0.0398  0.1412 -0.2821  56.8     0.78
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

"Tipping point" MNAR stress test

Treatment effects for different MNAR factors:

```
     tipping_factor    est     se      t  df Pr(>|t|)
[1,]          1.000 -2.717  0.927 -2.929 111   0.0041 **
[2,]          1.250 -2.401  0.933 -2.573 112   0.0114 *
[3,]          1.500 -2.085  0.939 -2.219 115   0.0284 *
[4,]          1.750 -1.769  0.947 -1.868 117   0.0643 .
[5,]          2.000 -1.453  0.956 -1.520 120   0.1311
[6,]          2.250 -1.137  0.966 -1.178 124   0.2412
[7,]          2.500 -0.821  0.977 -0.841 127   0.4018
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

"Tipping point" MNAR stress test

Imputed ADAS.m18 assuming under MAR

	id	female	age	group	ADAS11.m0	ADAS11.m6	ADAS11.m12	ADAS11.m18
1	4	1	75.4	active	22	24	29	29
2	8	0	75.0	active	12	14	15	15
3	10	1	70.0	placebo	11	19	13	27
4	17	1	80.1	active	21	18	20	21
5	29	1	66.2	placebo	13	11	17	13

Imputed ADAS.m18 assuming under MNAR (k=2.5, $k \times 4$ =10 ADAS points)

	id	female	age	group	ADAS11.m0	ADAS11.m6	ADAS11.m12	ADAS11.m18
1	4	1	75.4	active	22	24	29	39
2	8	0	75.0	active	12	14	15	25
3	10	1	70.0	placebo	11	19	13	27
4	17	1	80.1	active	21	18	20	31
5	29	1	66.2	placebo	13	11	17	13

Estimands, Estimators, & Estimates

- Estimand: Target of estimation
  e.g. the treatment effect in ITT population
- Estimator: Rules for estimation
  e.g. final visit contrast based on MMRM
- Estimate: The number derived from applying estimator to data
  e.g. -3.42 (SE=1.04) ADAS11 point difference between groups

Estimands: Efficacy vs Effectiveness

- Efficacy, de-jure, per-protocol: "an ideal treatment effect that could have been reached if all patients had fully adhered to the treatment"
- Effectiveness, de-facto, intention-to-treat: "the effect seen in practice"

Akacha, M., Bretz, F., Ruberg, S. (2017). Estimands in clinical trialsbroadening the perspective. *Statistics in Medicine*, 36(1), 5-19.

EMA draft guidline for AD

"[MMRM] tends to be less robust against a decreasing treatment effect difference after treatment discontinuation, which is one reason why in CNS indications the MMRM model often yields effect estimates close to those in the subgroup of patients who complete the study as planned. Therefore it is difficult to endorse the choice of the MMRM model as a routine approach to the primary analysis because of this concern that the results would tend to overestimate the true treatment effect."

Committee for Medicinal Products for Human Use (CHMP). Draft guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease and other dementias. European Medicines Agencies. 2016.

EMA draft guidline for AD

The EMA claim: MMRM provides estimates which are closer to *efficacy/per-protocol* than *effectiveness/ITT*.

- Theoretically at least, this claim is false if data after treatment discontinuation are included in the model and missingness is MAR
- Any method robust to MAR (e.g. mixed-effects model, multiple imputation) could have this same problem if data after treatment discontinuation are omitted
- Any alternative MNAR approach would necessarily make *untestable* assumptions about missing data
- Assessing the magnitude of this problem and potential alternative approaches is an active area of research ...
- Perhaps we should postone concerns about effectiveness until an efficacious treatment is identified

Further reading

- Mallinckrodt, C.H., et al. (2003). Assessing and interpreting treatment effects in longitudinal clinical trials with missing data. *Biological psychiatry*, 53(8), 754-760.
- Pinheiro, J.C., and Bates, D.M. (2000). Mixed-Effects Models in S and S-PLUS. Springer.
- Lu, K. (2010). On efficiency of constrained longitudinal data analysis versus longitudinal analysis of covariance. *Biometrics*, 66(3), 891-896.
- Donohue, M.C. and Aisen, P. S. (2012). Mixed model of repeated measures versus slope models in Alzheimer's disease clinical trials. *The journal of nutrition, health aging*. 16(4), 360-364.
- Little, R.J.A. and Rubin, D.B. (2014). Statistical analysis with missing data. John Wiley Sons.

Further reading

- van Buuren S, Groothuis-Oudshoorn K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 45, 1-67.
- Akacha, M., Bretz, F., Ruberg, S. (2017). Estimands in clinical trials-broadening the perspective. *Statistics in Medicine*, 36(1), 5-19.
- Committee for Medicinal Products for Human Use (CHMP). Draft guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease and other dementias. European Medicines Agencies. 2016.

Thank you!