Contemporary Issues in Clinical Trials Methods
Longitudinal Data Analysis Part I

Michael Donohue

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

July 14, 2017

USC ATRI

Learning objectives

- Define repeated measures/longitudinal data
- Identify study designs which give rise to longitudinal data
- Identify different analysis methods that are available
- Appreciate the complexities of proper analysis of longitudinal data
- Analyze data with ANCOVA and linear mixed-effects models

Further reading

- Diggle, P., Heagerty, P., Liang, K.-Y., Zeger, S. (2002). *Analysis of Longitudinal Data*. Oxford University Press
- Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics. New-York: Springer.
- Molenberghs, G. and Kenward, M.G. (2007). *Missing Data in Clinical Studies*. Wiley.

R packages

Follow along in R is not required, but we will use:

```
library(Hmisc)
library(tidyverse)
library(ggplot2)
library(nlme)
library(contrast)
```

If you do not have them installed already, you will need to download them from CRAN via:

```
install.packages(c("tidyverse", "Hmisc", "ggplot2", "nlme", "contrast"))
```

See `01LDA.R` for source code.

Repeated Measures/Longitudinal Data

Repeated measures are obtained when a response is measured **repeatedly** on a set of **observational units**

Observational Units:

- Subjects, patients, participants, ...
- Animals, plants, ...
- Clusters: universities, hospitals, clinical sites, towns, families, sets of twins, ...

Longitudinal data is a special (and popular) case of repeated measures in which the same observational unit is observed **over time**

Pre- post-treatment designs are a special (and popular) case of longitudinal data

Let's generate some data. . .

We simulate a randomized clinical trial with

- $n = 200$ mild to moderate dementia subjects per group
- Alzheimer's Disease Assessment Scale (ADAS-Cog13) assessed at 0, 6, 12, 18 months
- Weak effects for age and sex (based on ADNI pilot estimates)
- A treatment which slows ADAS-Cog13 progression by 25%

Let's generate some data. . .

First set some parameters:

```
# fixed effects parameters
Beta <- c('(Intercept)'=31.6, # mean ADAS at baseline
  female=-0.63, age=0.01,     # weak effects for sex and age
  month=0.44,                 # increase in ADAS per month in controls
  'month:active'=-0.11)       # relative slowing in active group

# random effects variance parameters
sigma_random_intercept <- 7.3
sigma_random_slope <- 0.45
sigma_residual <- 3.4

# other design parameters
months <- c(0, 6, 12, 18)
n <- 200 # per group
attrition_rate <- 0.05
```

Let's generate some data...

```
set.seed(20170701)

subjects <- data.frame(
  id = 1:(2*n),
  active = sample(c(rep(0,n), rep(1,n)), 2*n),
  female = sample(0:1, 2*n, replace=TRUE),
  age = rnorm(2*n, 75, 7.8),
  censor = rexp(2*n,rate=attrition_rate),
  ran.intercept = rnorm(2*n, sd=sigma_random_intercept),
  ran.slope     = rnorm(2*n, sd=sigma_random_slope))

trial <- right_join(subjects,
  expand.grid(id = 1:(2*n), month=months)) %>%
  mutate(
    residual = rnorm(2*n*length(months), sd=sigma_residual),
    group = factor(active, 0:1, c('placebo', 'active')),
    missing = ifelse(month>censor, 1, 0)) %>%
  arrange(id, month)
```

Let's generate some data...

```
trial$ADAS13 <- round(
  model.matrix(~ female+age+month+month:active, data = trial)[, names(Beta)] %*%
  Beta +
  with(trial, ran.intercept + ran.slope*month + residual),
  digits = 0
)[,1]
```

```
  id active female  age  censor ran.intercept ran.slope month residual   group missing ADAS13
1  1      0      1 81.0   12.8          -7.99     0.718     0     3.63 placebo       0     27
2  1      0      1 81.0   12.8          -7.99     0.718     6    -4.65 placebo       0     26
3  1      0      1 81.0   12.8          -7.99     0.718    12     4.83 placebo       0     43
4  1      0      1 81.0   12.8          -7.99     0.718    18     2.87 placebo       1     48
5  2      1      0 75.6   32.0         -12.09    -0.497     0     2.31  active       0     23
6  2      1      0 75.6   32.0         -12.09    -0.497     6    -3.63  active       0     16
```

Let's generate some data. . .

```r
# filter out the missing observations
trial_obs <- filter(trial, !missing)

# transfrom data from long to wide
trial_wide <- trial_obs %>%
  select(id, month, female, age, active, group, ADAS13) %>%
  mutate(month = paste0('ADAS13.m', month)) %>%
  spread(month, ADAS13)

# data for MMRM
trial_mmrm <- right_join(
  select(trial_wide, id, ADAS13.m0),
  filter(trial_obs, month>0))
```

```
  id active female  age  censor ran.intercept ran.slope month residual   group missing ADAS13
1  1      0      1 81.0    12.8         -7.99     0.718     0     3.63 placebo       0     27
2  1      0      1 81.0    12.8         -7.99     0.718     6    -4.65 placebo       0     26
3  1      0      1 81.0    12.8         -7.99     0.718    12     4.83 placebo       0     43
4  2      1      0 75.6    32.0        -12.09    -0.497     0     2.31  active       0     23
5  2      1      0 75.6    32.0        -12.09    -0.497     6    -3.63  active       0     16
6  2      1      0 75.6    32.0        -12.09    -0.497    12    -3.45  active       0     15
```

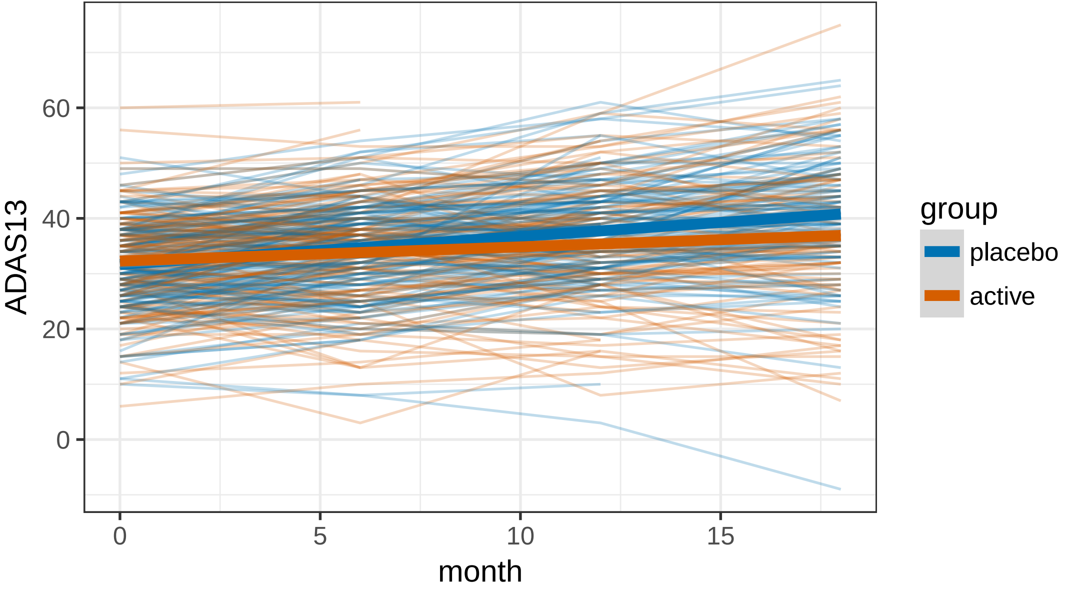
Baseline characteristics

Table: Descriptive Statistics by group

	placebo $N = 200$	active $N = 200$
female	49% (98)	55% (109)
age	68.36 75.19 79.85 (74.58 ± 8.43)	70.33 75.59 80.80 (75.69 ± 7.78)
ADAS13	26.00 31.00 38.00 (31.75 ± 7.85)	26.00 33.00 38.00 (32.22 ± 8.34)

a b c represent the lower quartile a, the median b, and the upper quartile c for continuous variables. $x \pm s$ represents $\bar{X} \pm 1$ SD.Numbers after percents are frequencies.

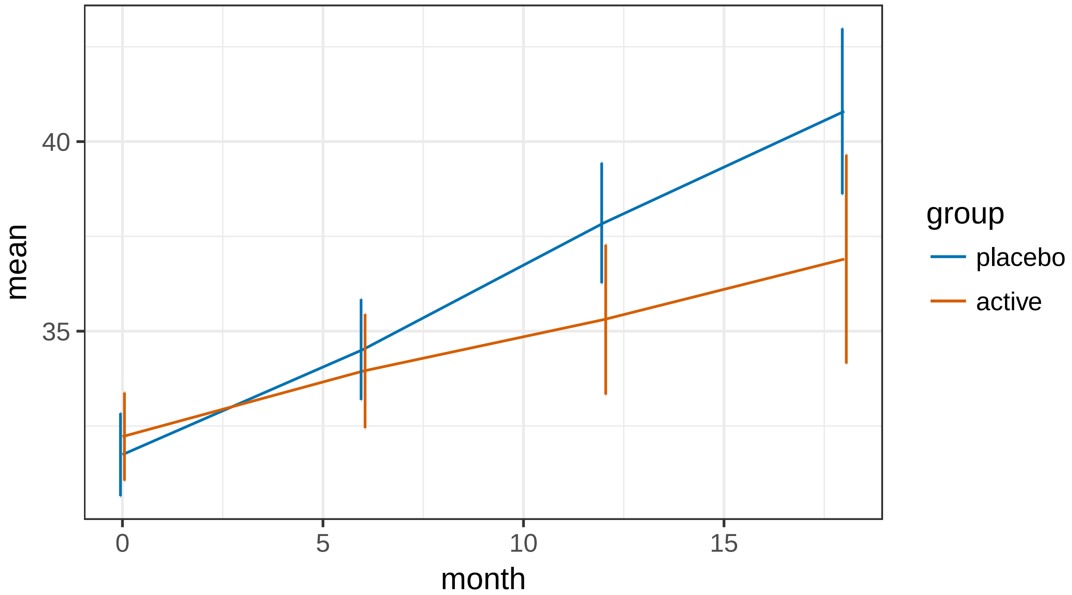
Spaghetti plot



Basic longitudinal summaries of ADAS13

group	month	n	mean	sd	lower95	upper95	min	max
placebo	0	200	31.7	7.85	30.7	32.8	10	51
placebo	6	155	34.5	8.38	33.2	35.8	8	54
placebo	12	130	37.9	9.15	36.3	39.4	3	61
placebo	18	104	40.8	11.25	38.6	43.0	-9	65
active	0	200	32.2	8.34	31.1	33.4	6	60
active	6	155	33.9	9.48	32.4	35.5	3	61
active	12	118	35.3	10.84	33.3	37.3	8	59
active	18	91	36.9	13.22	34.1	39.7	7	75

Mean plot



Two sample t-test of group difference at month 18 (completers analysis)

- Difference between group means is 40.798 - 36.901 = 3.897
- (pooled) standard deviation is 12.207
- $t = \frac{3.897}{12.207\sqrt{\frac{1}{104}+\frac{1}{91}}} = 2.224$
- 104 + 91 - 2 = 193 "degrees of freedom"

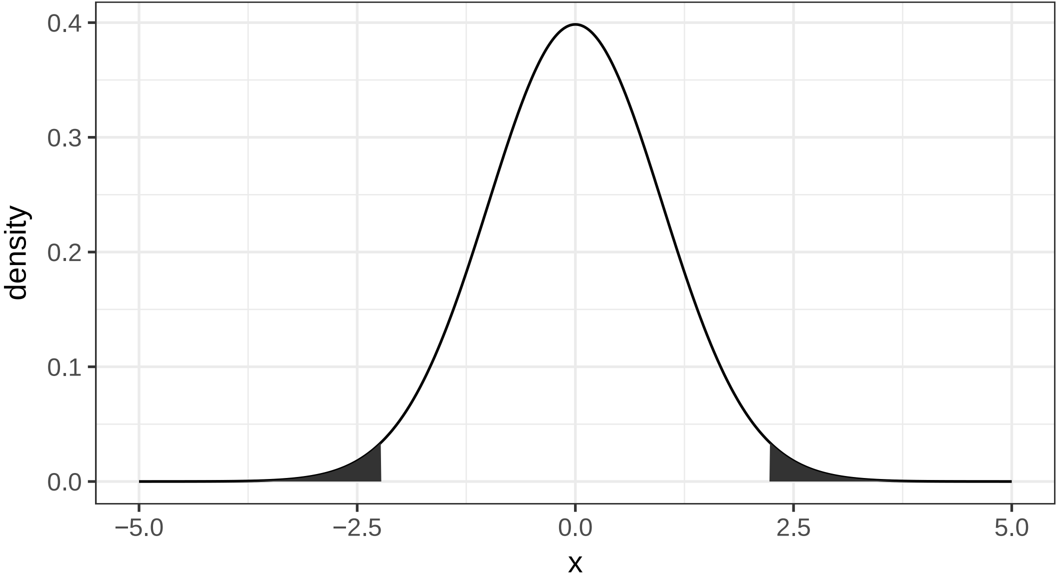
t-test

```
Two Sample t-test

data:  ADAS13 by group
t = 2.22, df = 193, p-value = 0.0273
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.441 7.353
sample estimates:
mean in group placebo  mean in group active
              40.7981               36.9011
```

The t_{193}-distribution

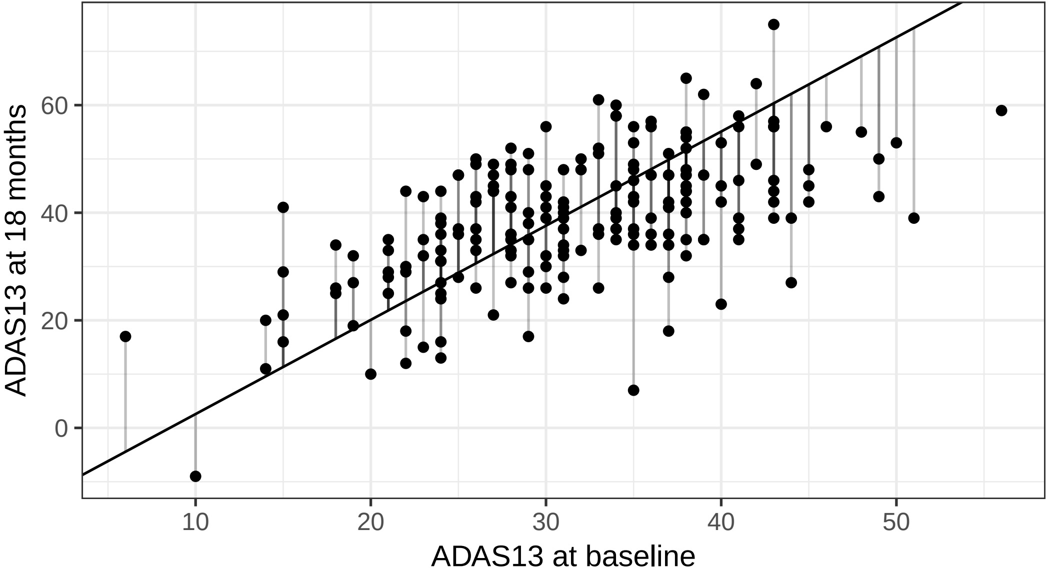
p-value is area under curve for $|x| > 2.224$, the value of the test statistic in this case.



Regression analysis

- "Regression" generally refers to a relationship between variables that is estimated by data
- "Ordinary Least Squares" regression, for example, describes a linear relationship between two continuous variables that is estimated by the line that minimizes the sum of squared "residuals"
- "Residuals" are the differences between observations and values predicted by the regression

Ordinary Least Squares



Other types of regression

- "General linear models" can add multiple *covariates/predictors*
- "Generalized linear models" can accomodate other types of *outcome/response* variables (e.g. logistic regression can accomodate binary outcome variables)
- "Mixed-effects models" mix *random effects* with the standard *fixed effects* to account for complex correlation structures

All regression models share the common theme of estimating the best fit relationship between *outcome/response* variables and *covariates/predictors*

ANalysis of COVAriance (ANCOVA) for "*pre-post*" data

- Very common for two groups, and *one* post- assessment
- Y_{i1}: baseline or pre- observation
- Y_{i2}: followup or post- observation
- Trt_i: treatment group indicator (e.g. 1 if active, 0 if placebo)
- ANCOVA I: $Y_{i2} = \beta_0 + \mathrm{Trt}_i\beta_1 + Y_{i1}\beta_2 + \varepsilon_i$
    - Includes effect for the outcome at baseline
    - β_1 is the effect of interest
- ANCOVA II: $Y_{i2} = \beta_0 + \mathrm{Trt}_i\beta_1 + Y^*_{i1}\beta_2 + \mathrm{Trt}_i Y^*_{i1}\beta_3 + \varepsilon_i$
    - Includes additional effect for the interaction between treatment and outcome at baseline
    - β_1 is the effect of interest
    - Need to *mean center* baseline covariates: $Y^*_{i1} = Y_{i1} - \bar{Y}_{\cdot 0}$

Yang & Tsiatis (2001). Efficiency Study of Estimators for a Treatment Effect in a Pretest-Posttest Trial. *The Am. Statistician*, 55(4) 314-321

ANCOVA I for effect of treatment on ADAS13 at 18 months

```
Call:
lm(formula = ADAS13.m18 ~ active + ADAS13.m0, data = trial_wide)

Residuals:
   Min     1Q Median     3Q    Max
-32.73  -6.05  -0.45   6.45  28.45

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  13.8803     2.8631    4.85  2.6e-06 ***
active       -3.9777     1.4247   -2.79   0.0058 **
ADAS13.m0     0.8522     0.0852   10.00  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.92 on 192 degrees of freedom
  (205 observations deleted due to missingness)
Multiple R-squared:  0.359,Adjusted R-squared:  0.352
F-statistic: 53.7 on 2 and 192 DF,  p-value: <2e-16
```

ANCOVA II for effect of treatment on ADAS13 at 18 months

```
Call:
lm(formula = ADAS13.m18 ~ active * center(ADAS13.m0), data = trial_wide)

Residuals:
   Min     1Q Median     3Q    Max
-32.61  -6.11  -0.43   6.30  28.87

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)               41.1511     0.9765   42.14  < 2e-16 ***
active                    -4.0045     1.4289   -2.80   0.0056 **
center(ADAS13.m0)          0.8916     0.1227    7.27  9.2e-12 ***
active:center(ADAS13.m0)  -0.0765     0.1709   -0.45   0.6549
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.95 on 191 degrees of freedom
  (205 observations deleted due to missingness)
Multiple R-squared:  0.359,Adjusted R-squared:  0.349
F-statistic: 35.7 on 3 and 191 DF,  p-value: <2e-16
```

ANCOVA II with more covariates

```
Call:
lm(formula = ADAS13.m18 ~ active * center(ADAS13.m0) + female +
    age, data = trial_wide)

Residuals:
   Min     1Q Median     3Q    Max
-31.85  -5.96  -0.78   6.84  29.31

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)               47.4017     6.6519    7.13  2.1e-11 ***
active                    -4.0012     1.4298   -2.80   0.0057 **
center(ADAS13.m0)          0.9013     0.1228    7.34  6.3e-12 ***
female                     1.4720     1.4430    1.02   0.3090
age                       -0.0931     0.0866   -1.07   0.2840
active:center(ADAS13.m0)  -0.0754     0.1712   -0.44   0.6601
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.94 on 189 degrees of freedom
  (205 observations deleted due to missingness)
Multiple R-squared:  0.367,Adjusted R-squared:  0.35
F-statistic: 21.9 on 5 and 189 DF,  p-value: <2e-16
```

ANCOVA summary

- Ubiquitous, simple, powerful framework
- Inherently a **complete case** analysis!
- With missing data, not intention-to-treat (ITT) analysis
- Does not make use of **incomplete cases!**
- Might be biased and/or inefficient (low power) with missing data

Two-stage models

- *Subject-specific* longitudinal profiles can often be modeled with simple linear regression
- This leads to the 2-stage model:
  - Stage 1: Linear regression model for each subject separately
  - Stage 2: Model subject-specific regression coefficients with covariates of interest
- However, this is **NOT** a recommend analysis approach, but rather a means to introduce mixed-effect models.

Two-stage model example

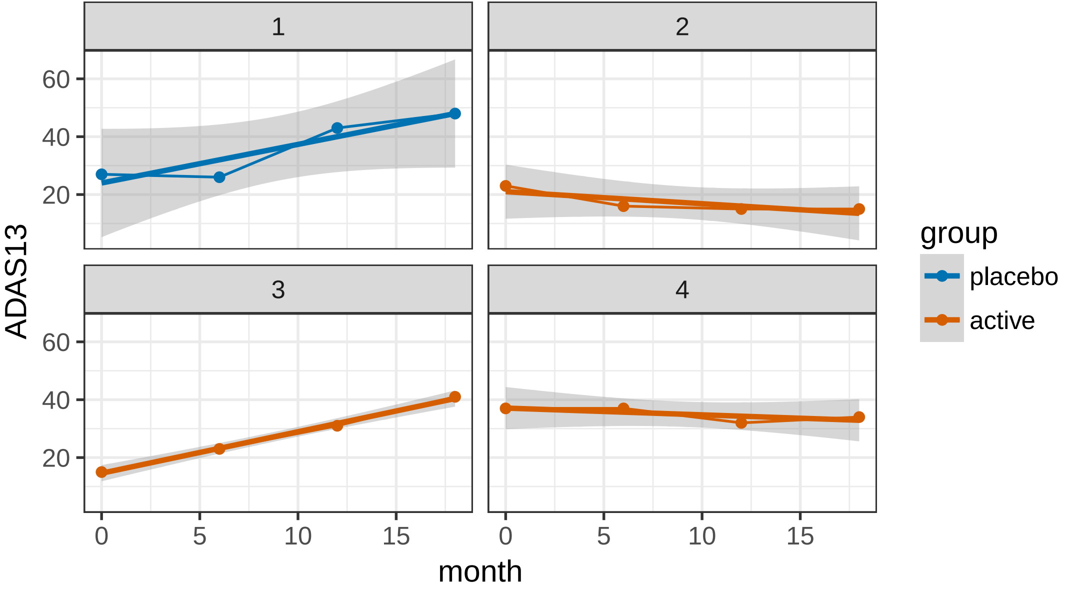
- Stage 1:

$$Y_{ij} = \beta_{0i} + t_{ij}\beta_{1i} + \varepsilon_i \tag{1}$$

  for subject i at time t_{ij}
- Provides estimates of subject-specific intercepts, $\hat{\beta}_{0i}$ and slopes $\hat{\beta}_{0i}$
- $\varepsilon_i \sim N(0, \sigma_i^2 I_{n_i})$ estimates *within*-subject variability
- *Between*-subject variability can now be modeled by treating $\hat{\beta}_i$ as "response variables"
- Stage 2:

$$\hat{\beta}_{1i} = X_i\beta + \varepsilon_i' \tag{2}$$

Stage 1 models of simulated trial



Stage 1 model of simulated trial

```
  id beta.(Intercept) beta.month sigma active   group  age female
1  1             24.0      1.333 7.348      0 placebo 81.0      1
2  2             21.0     -0.417 2.598      1  active 75.6      0
3  3             14.6      1.433 0.775      1  active 70.0      0
4  4             37.1     -0.233 2.025      1  active 71.5      0
5  5             41.0      0.667   NaN      1  active 84.1      1
6  6             39.0         NA   NaN      1  active 76.2      1
```

Stage 1 model of simulated trial

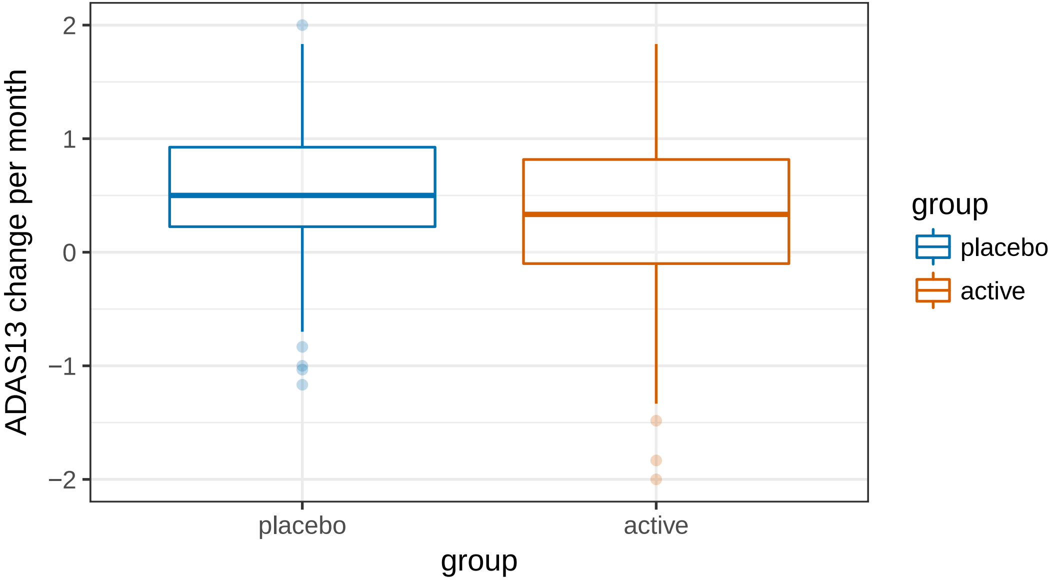
```
Call:
lm(formula = ADAS13 ~ month, data = trial_obs, subset = id ==
    1)

Residuals:
 1  2  3
 3 -6  3

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   24.000      6.708    3.58     0.17
month          1.333      0.866    1.54     0.37

Residual standard error: 7.35 on 1 degrees of freedom
Multiple R-squared:  0.703,Adjusted R-squared:  0.407
F-statistic: 2.37 on 1 and 1 DF,  p-value: 0.367
```

Stage 2 model of simulated trial



Stage 2 model of simulated trial

```
Call:
lm(formula = beta.month ~ female + age + active, data = trial_stage1)

Residuals:
    Min      1Q  Median      3Q     Max
-2.3945 -0.3799 -0.0128  0.4089  1.6386

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.56106    0.34781    1.61   0.1077
female       0.20069    0.07400    2.71   0.0071 **
age         -0.00205    0.00456   -0.45   0.6525
active      -0.20108    0.07390   -2.72   0.0069 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.648 on 306 degrees of freedom
  (90 observations deleted due to missingness)
Multiple R-squared:  0.0464,Adjusted R-squared:  0.0371
F-statistic: 4.96 on 3 and 306 DF,  p-value: 0.00223
```

Two-stage models

- In contrast to ANCOVA and t-test, two-stage models allow all randomized subject with at least one followup to be included into analysis ("modified intention-to-treat")
- However, second stage models ignore the variability/uncertainty of the slope estimates from the first stage
- This means that p-values from second stage might be smaller than they should be and Type I error could be inflated

Linear mixed-effects models

Linear mixed-effects models provide a cleaner, more efficient, and more accurate one-step alternative to two-stage models

$$\left.\begin{aligned} \text{Stage 1: } Y_{ij} &= \beta_{0i} + t_{ij}\beta_{1i} + \varepsilon_i \\ \text{Stage 2: } \hat{\beta}_{1i} &= X_i\beta + \varepsilon_i' \end{aligned}\right\} \rightarrow Y_{ij} = X_i\beta + b_{0i} + t_{ij}b_{1i} + \varepsilon_i$$

β	population level "fixed effects"
$b_i \sim N(0, D)$	subject-specific "random effects" for sujbect i
$\varepsilon_i \sim N(0, \Sigma_i)$	vector of "residuals" for subject i
D, Σ_i	"variance components"

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

Linear mixed-effects models of simulated trial

```
Linear mixed-effects model fit by REML
 Data: trial_obs
   AIC  BIC logLik
  7565 7600  -3775

Random effects:
 Formula: ~month | id
 Structure: General positive-definite, Log-Cholesky parametrization
            StdDev Corr
(Intercept) 7.370  (Intr)
month       0.478  -0.008
Residual    3.529

Fixed effects: ADAS13 ~ month + month:active
             Value Std.Error  DF t-value p-value
(Intercept)   32.0     0.401 751    79.9  0.0000
month          0.5     0.047 751    10.9  0.0000
month:active  -0.2     0.067 751    -3.1  0.0021
 Correlation:
             (Intr) month
month        -0.107
month:active -0.005 -0.691

Standardized Within-Group Residuals:
    Min      Q1     Med      Q3     Max
-2.6244 -0.4800  0.0047  0.4731  3.2822

Number of Observations: 1153
Number of Groups: 400
```

LME model with additional covariates

```
Linear mixed-effects model fit by REML
 Data: trial_obs
   AIC  BIC logLik
  7565 7610  -3773

Random effects:
 Formula: ~month | id
 Structure: General positive-definite, Log-Cholesky parametrization
            StdDev Corr
(Intercept) 7.299  (Intr)
month       0.478  0.01
Residual    3.528

Fixed effects: ADAS13 ~ age + female + month + month:active
             Value Std.Error  DF t-value p-value
(Intercept)   37.0      3.70 751    9.99  0.0000
age           -0.1      0.05 397   -1.07  0.2841
female        -2.0      0.79 397   -2.49  0.0131
month          0.5      0.05 751   10.87  0.0000
month:active  -0.2      0.07 751   -3.05  0.0023
 Correlation:
             (Intr) age    female month
age          -0.988
female       -0.128  0.018
month        -0.019  0.008 -0.001
month:active  0.012 -0.012 -0.005 -0.692

Standardized Within-Group Residuals:
      Min        Q1       Med        Q3       Max
-2.605750 -0.471148 -0.000855  0.469398  3.272736

Number of Observations: 1153
Number of Groups: 400
```

Linear mixed-effects models (R code)

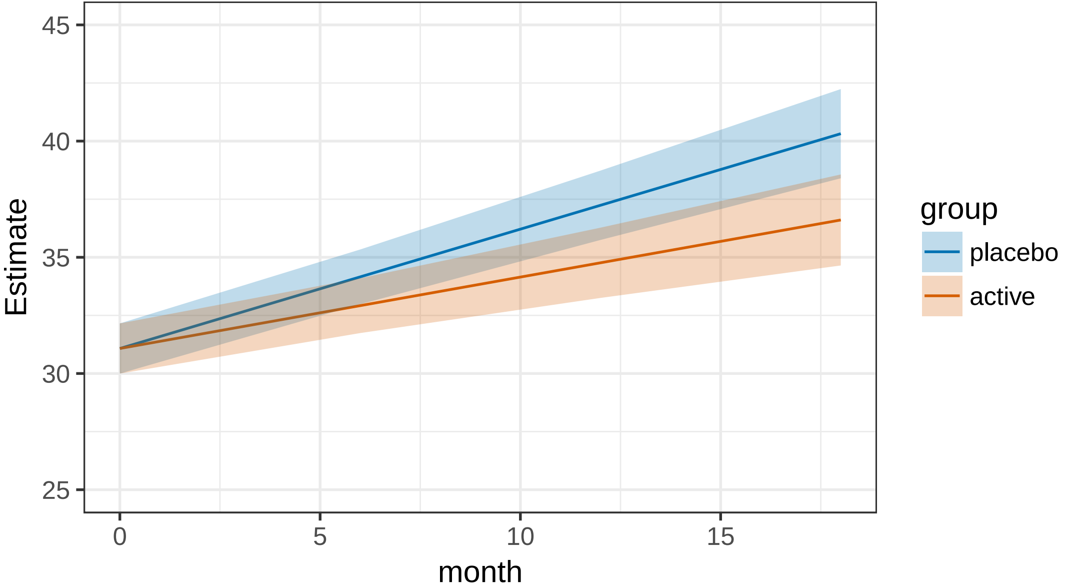
```
lme(ADAS13 ~ month + month:active, data = trial_obs,
  random = ~month|id)

lme(ADAS13 ~ age + female + month + month:active, data = trial_obs,
  random = ~month|id)
```

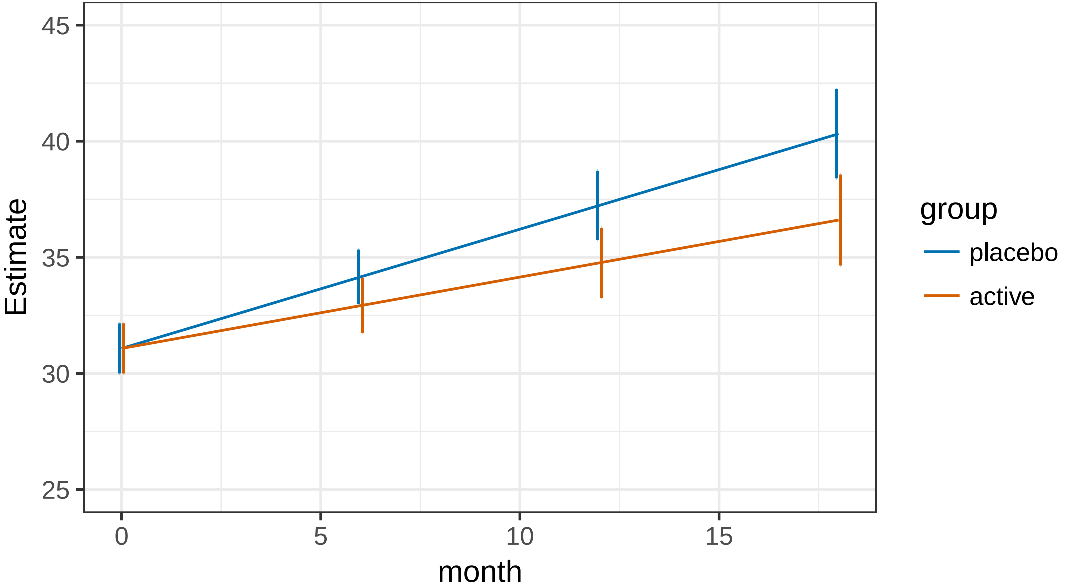
Mean profiles

```
active month.active month  age Estimate Lower Upper   group
     1            0     0 75.1     31.1  30.0  32.2  active
     1            6     6 75.1     32.9  31.7  34.1  active
     1           12    12 75.1     34.8  33.3  36.3  active
     1           18    18 75.1     36.6  34.6  38.6  active
     0            0     0 75.1     31.1  30.0  32.2 placebo
     0            0     6 75.1     34.2  33.0  35.3 placebo
     0            0    12 75.1     37.2  35.7  38.7 placebo
     0            0    18 75.1     40.3  38.4  42.2 placebo
```

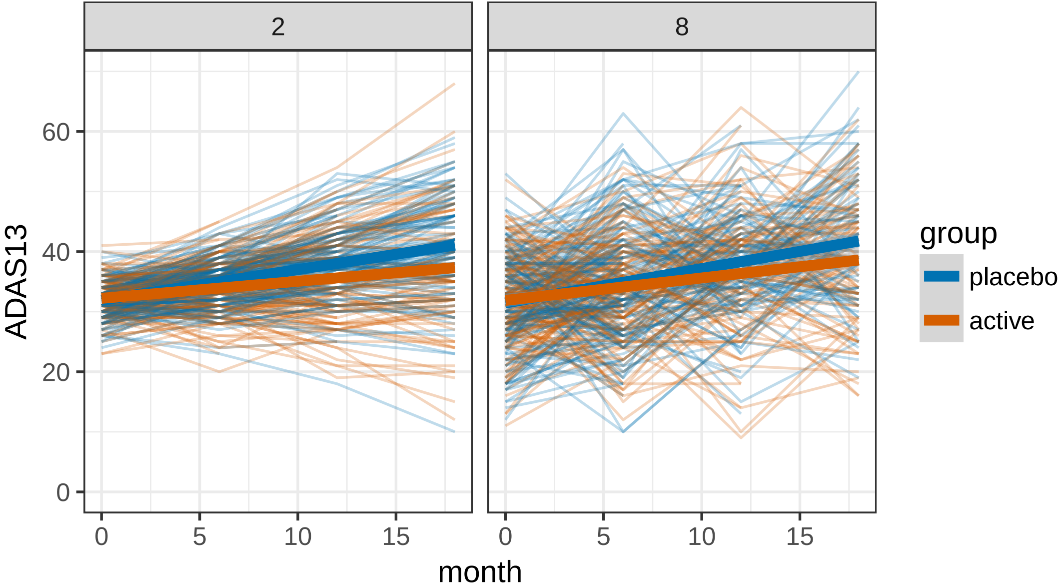
Plotting profiles (shaded CIs)



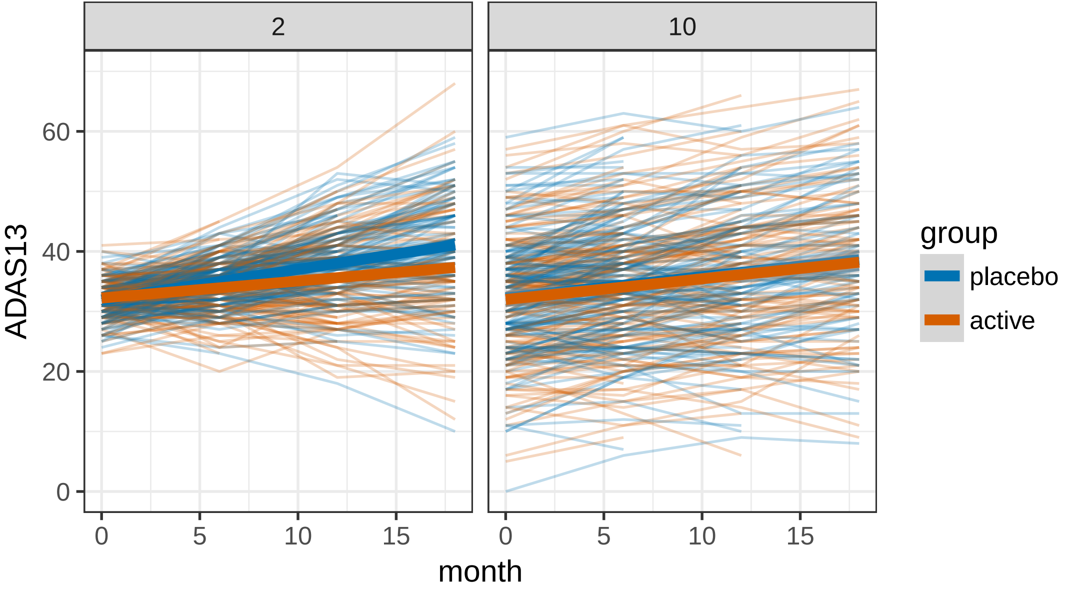
Plotting profiles (error bar CIs)



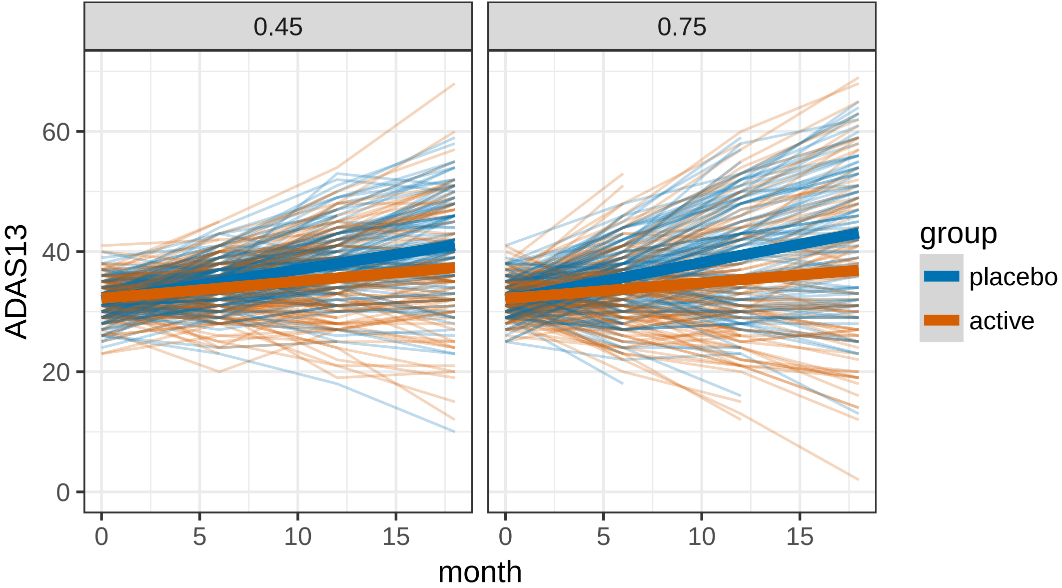
Mixed effect models: standard deviation of residuals



Mixed effect models: standard deviation of random intercepts



Mixed effect models: standard deviation of random slopes



Random intercepts model

If we drop the *random slope* term, what remains is called a *random intercepts* model

$$Y_{ij} = X_i\beta + b_{0i} + t_{ij}b_{1i} + \varepsilon_i$$

Random intercepts model

```
Linear mixed-effects model fit by REML
 Data: trial_obs
   AIC  BIC logLik
  7781 7807  -3886

Random effects:
 Formula: ~1 | id
        (Intercept) Residual
StdDev:        7.83     4.95

Fixed effects: ADAS13 ~ month + month:active
             Value Std.Error  DF t-value p-value
(Intercept)   32.0     0.448 751    71.4       0
month          0.5     0.033 751    15.7       0
month:active  -0.2     0.046 751    -4.7       0
 Correlation:
             (Intr) month
month        -0.236
month:active -0.006 -0.668

Standardized Within-Group Residuals:
     Min       Q1      Med       Q3      Max
-4.21389 -0.51428  0.00898  0.51771  3.55703

Number of Observations: 1153
Number of Groups: 400
```

Random intercepts model vs model with random slopes

```
anova(fit_lme_int, fit_lme)

            Model df  AIC  BIC logLik   Test L.Ratio p-value
fit_lme_int     1  5 7781 7807  -3886
fit_lme         2  7 7565 7600  -3775 1 vs 2     220  <.0001
```

The model with random slopes is preferred