

Basic Statistic for health researchers

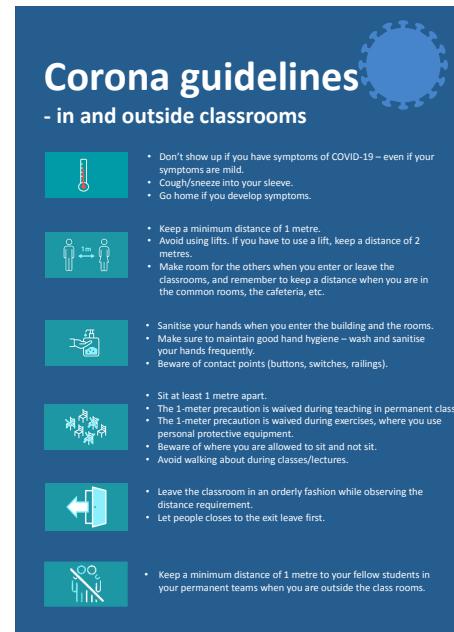
Lecture 8: repeated measurements

Brice Ozenne - email: broz@sund.ku.dk

¹ Section of Biostatistics, Department of Public Health, University of Copenhagen

² Neurobiology Research Unit, University Hospital of Copenhagen, Rigshospitalet.

23-11-2020



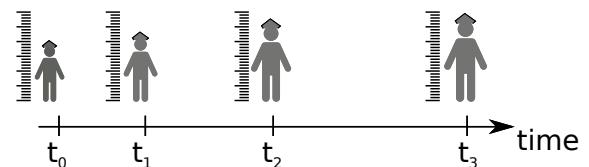
0 / 44

General concepts

What are repeated measurements?

Measurement of a variable at different occasions on an experimental unit:

- typically **same** type of **measurement** on the **same patient** at **different timepoints**.



Introduction
●○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○

Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○○○○

Introduction
●○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○○

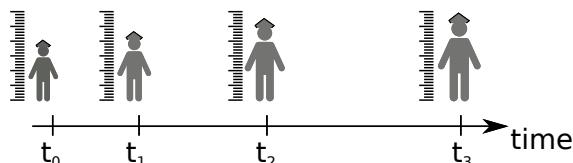
Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○○○○

What are repeated measurements?

Measurement of a variable at different occasions on an experimental unit:

- typically **same** type of **measurement** on the **same patient** at **different timepoints**.



To make things simple, we will focus on:

- comparing two treatment groups
- at fixed timepoints
- no missing data

Basic Statistics for health researchers - L8: Repeated measurements

2 / 44

Introduction
●○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○

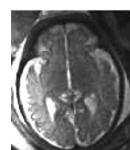
Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○○○○

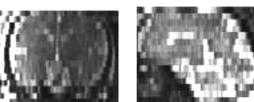
What are repeated measurements?

Measurement of a variable at different occasions on an experimental unit:

- could also be **different** ways of **measuring** the same quantity on the **same patient** at the **same timpoint and location**.

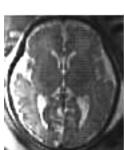


Gold standard

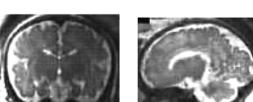


(a) Original (Axial) (b) Original (Coronal) (c) Original (Sagittal)

VS.



High resolution



(d) HR (Axial) (e) HR (Coronal) (f) HR (Sagittal)

(brain pictures from Van Reeth et al. (2012))

Basic Statistics for health researchers - L8: Repeated measurements

2 / 44

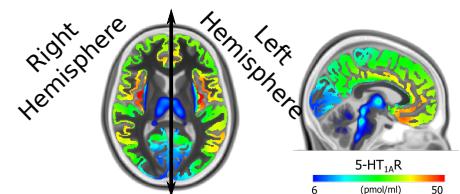
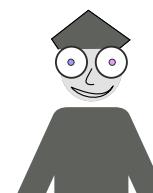
Basic Statistics for health researchers - L8: Repeated measurements

3 / 44

What are repeated measurements?

Measurement of a variable at different occasions on an experimental unit:

- could also be the **same** type of **measurement** on the **same patient** at **different locations**.



(brain picture from Beliveau et al. (2017))

Basic Statistics for health researchers - L8: Repeated measurements

2 / 44

Introduction
●○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○○

Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○○○○

Why performing repeated measurements? (1/4)

Introduction
○○○○○
○○○

Pre-post study
●○○○○○○
○○○○○
○○○○

Multivariate model
○○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○○○○

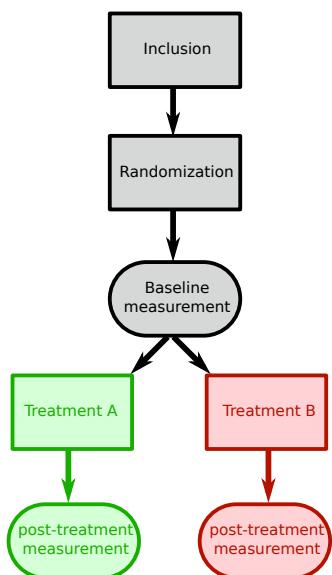
Introduction
○○○○○
○○○

Pre-post study
○●○○○○○○
○○○○○
○○○○

Multivariate model
○○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○○○○

Design: pre-post study



Basic Statistics for health researchers - L8: Repeated measurements

Introduction
○○○○○
○○○

Pre-post study
○●○○○○○
○○○○○
○○○○

Multivariate model
○○○○○
○○
○○○○○

11 / 44

Age-Related Macular Degeneration (ARMD) Trial:

- clinical trial comparing interferon- α and placebo
- baseline + week 4 measurements

Main outcome:

- change in vision between week 4 and baseline

Illustration: armd dataset from the *nlmeU* package

```
dW.pp <- read.table("prepost.txt")
head(dW.pp)
```

	subject	treatment	lesion	visual0	visual4
1	1	Active	3	59	55
2	2	Active	1	65	70
3	3	Placebo	4	40	40
4	4	Placebo	2	67	64
6	6	Active	3	59	53
7	7	Placebo	1	64	68

Basic Statistics for health researchers - L8: Repeated measurements

13 / 44

Basic Statistics for health researchers - L8: Repeated measurements

Basic Statistics for health researchers - L8: Repeated measurements

Introduction
○○○○○
○○○

Pre-post study
○●○○○○○
○○○○○
○○○○

Multivariate model
○○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○○○○

12 / 44

Moving to the long format (1/3)

subject	group	outcome	
		(time = 0)	(time = 4)
1	A	59	55
2	A	65	70
3	P	40	40
..

Long

14 / 44

Moving to the long format (1/3)

Wide

subject	group	outcome	
		(time = 0)	(time = 4)
1	A	59	55
2	A	65	70
3	P	40	40
..

Long

subject	group	time	outcome
		0	
		4	
		0	
		4	
		..	

Wide

subject	group	(time = 0)	(time = 4)
1	A	59	55
2	A	65	70
3	P	40	40
..

Long

subject	group	time	outcome
1	A	0	59
1	A	4	55
?	?	0	?
?	?	4	?
		0	
		4	
		..	

Moving to the long format (2/3)

```
## reshape to long format
dL.pp <- reshape2::melt(dW.pp,
  id.var = c("subject", "treatment", "lesion"),
  measure.vars = c("visual0", "visual4"),
  variable.name = "week",
  value.name = "visual")
```

```
## re-order dataset
dL.pp <- dL.pp[order(dL.pp$subject),]
## convert week as factor with appropriate values
dL.pp$week <- factor(dL.pp$week,
  level = c("visual0", "visual4"),
  labels = c(0,4))
## remove row names
rownames(dL.pp) <- NULL
```

Moving to the long format (3/3)

```
dW.pp[1:3,] # Wide format: 1 line = 1 subject
```

subject	treatment	lesion	visual0	visual4
1	1	Active	3	59
2	2	Active	1	65
3	3	Placebo	4	40

```
head(dL.pp) # Long format: 1 line = 1 measurement
```

subject	treatment	lesion	week	visual
1	1	Active	3	0
2	1	Active	3	4
3	2	Active	1	0
4	2	Active	1	4
5	3	Placebo	4	0
6	3	Placebo	4	4

Introduction
○○○○○
○○○

Pre-post study
○○○○○●○
○○○○

Multivariate model
○○○○
○○○○○

Conclusion
○○○
○○○○○○○○○

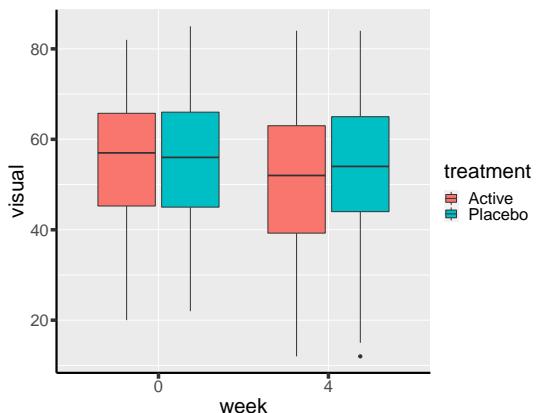
Introduction
○○○○○
○○○

Multivariate model
○○○○
○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

Graphical display: boxplot

```
gg <- ggplot(data = dL.pp,  
               mapping = aes(x=week, y=visual, fill=treatment))  
gg + geom_boxplot()
```



Basic Statistics for health researchers - L8: Repeated measurements

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○
●○○○○

Multivariate model
○○○○
○○○○○○○○

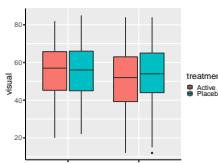
Conclusion
○○○
○○○○○○○○○

17 / 44

Testing the treatment effect - exercise!

Assuming normally distributed "visual" values:

- null hypothesis relative to the treatment effect?
- how can you summarize the data?
- re-phrase the null hypothesis in mathematical terms

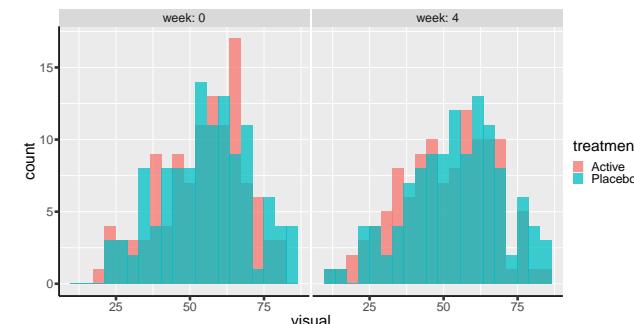


Basic Statistics for health researchers - L8: Repeated measurements

19 / 44

Graphical display: histogram

```
library(ggplot2)  
gg <- ggplot(data = dL.pp,  
               mapping = aes(x = visual, fill=treatment))  
gg <- gg + facet_wrap(~week, labeller = label_both)  
gg + geom_histogram(position = "identity",  
                     alpha = 0.75, bins = 20)
```



Basic Statistics for health researchers - L8: Repeated measurements

Introduction
○○○○○
○○○

Multivariate model
○○○○
○○○○○○○○

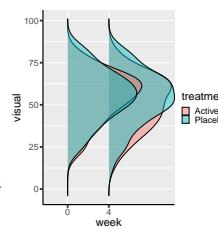
Conclusion
○○○
○○○○○○○○○○

18 / 44

Testing the treatment effect - exercise!

Assuming normally distributed "visual" values:

- null hypothesis relative to the treatment effect?
- how can you summarize the data?
- re-phrase the null hypothesis in mathematical terms



Basic Statistics for health researchers - L8: Repeated measurements

19 / 44

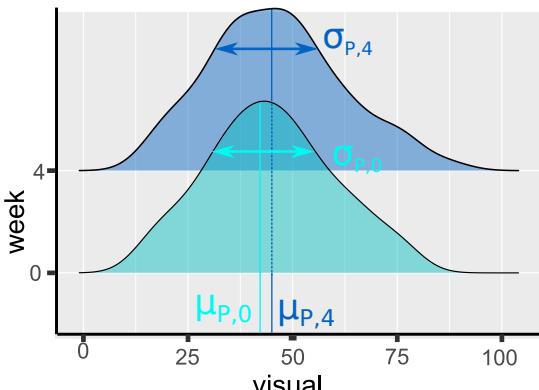
Introduction
○○○○○
○○○Pre-post study
○○○○○○○○
○○●○○○
○○○○Multivariate model
○○○○○
○○
○○○○○Conclusion
○○○
○○○○○○○○○○Pre-post study
○○○○○○○○
○○●○○○
○○○○○Multivariate model
○○○○○
○○
○○○○○Conclusion
○○○
○○○○○○○○○○

Testing the treatment effect - solution!

In plain english:

H_0 : same expected change in "visual" value in the Active and Placebo arm

Summarize the data:



Basic Statistics for health researchers - L8: Repeated measurements

Introduction
○○○○○
○○○Pre-post study
○○○○○○○○
○○●○○○
○○○○Multivariate model
○○○○○
○○
○○○○○

20 / 44

Basic Statistics for health researchers - L8: Repeated measurements

Pre-post study
○○○○○○○○
○○●○○○
○○○○○Multivariate model
○○○○○
○○
○○○○○Conclusion
○○○
○○○○○○○○○○

Estimating the treatment effect

$H_0: \mu_{A,4} - \mu_{A,0} = \mu_{P,4} - \mu_{P,0}$

Estimation: using the R function **mean**

- in each subgroups, subsetting (manually) the dataset

```
dW.ppA <- dW.pp[dW.pp$treatment=="Active",]
dW.ppP <- dW.pp[dW.pp$treatment=="Placebo",]
c(mu_A.0=mean(dW.ppA[, "visual0"]),
  mu_P.0=mean(dW.ppP[, "visual0"]))
```

mu_A.0 mu_P.0

54.42105 55.26496

- in all subgroups at once, using the long format

Basic Statistics for health researchers - L8: Repeated measurements

21 / 44

Basic Statistics for health researchers - L8: Repeated measurements

22 / 44

Testing the treatment effect - solution!

In plain english:

H_0 : same expected change in "visual" value in the Active and Placebo arm

Summarize the data:

- average value_{group,time}: $\mu_{A,0}, \mu_{A,4}, \mu_{P,0}, \mu_{P,4}$
- variance_{group,time}: $\sigma_{A,0}^2, \sigma_{A,4}^2, \sigma_{P,0}^2, \sigma_{P,4}^2$

Mathematically:

$H_0: \mu_{A,4} - \mu_{A,0} = \mu_{P,4} - \mu_{P,0}$

Basic Statistics for health researchers - L8: Repeated measurements

Introduction
○○○○○
○○○Pre-post study
○○○○○○○○
○○●○○○
○○○○Multivariate model
○○○○○
○○
○○○○○

20 / 44

Basic Statistics for health researchers - L8: Repeated measurements

Pre-post study
○○○○○○○○
○○●○○○
○○○○○Multivariate model
○○○○○
○○
○○○○○Conclusion
○○○
○○○○○○○○○○

Estimating summary statistics "at once" (1/2)

Example: average outcome value **at each timepoint**

Long			
subject	group	time	outcome
1	A	0	59
1	A	4	55
2	A	0	65
2	A	4	70
3	P	0	40
3	P	4	40
..

Estimating summary statistics "at once" (1/2)

Example: average outcome value **at each timepoint**

Sort

Long			
subject	group	time	outcome
1	A	0	59
1	A	4	55
2	A	0	65
2	A	4	70
3	P	0	40
3	P	4	40
..

Long			
subject	group	time	outcome
1	A	0	59
2	A	0	65
3	P	0	40
..	..	0
1	A	4	55
2	A	4	70
3	P	4	40
..	..	4

Basic Statistics for health researchers - L8: Repeated measurements

22 / 44

Pre-post study	Multivariate model	Conclusion
○○○○○○○○	○○○○	○○○
○○○●○	○○	○○○○○○○○
○○○○	○○○○○○	

Estimating summary statistics "at once" (1/2)

Example: average outcome value **at each timepoint**

Split 1				Summarize
subject	group	time	outcome	
1	A	0	59	
2	A	0	65	
3	P	0	40	
..	..	0	

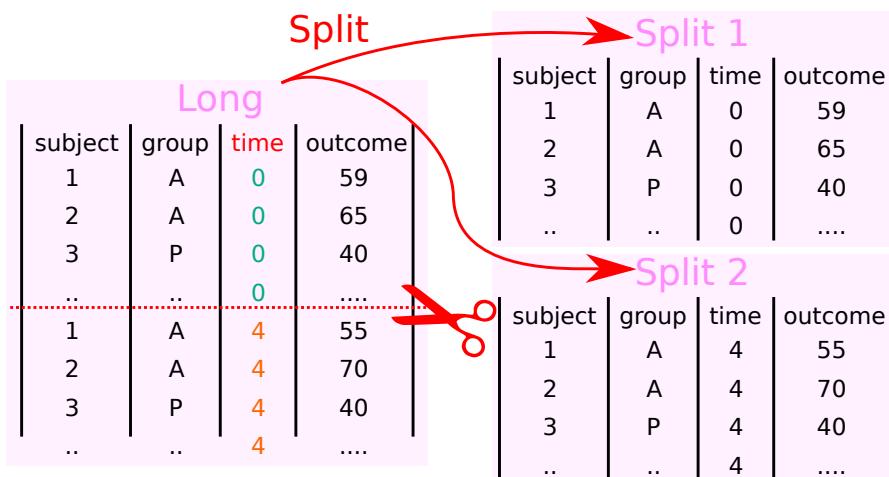
Split 2				mean
subject	group	time	outcome	
1	A	4	55	59
2	A	4	70	65
3	P	4	40	40
..	..	4

Basic Statistics for health researchers - L8: Repeated measurements

22 / 44

Estimating summary statistics "at once" (1/2)

Example: average outcome value **at each timepoint**



Basic Statistics for health researchers - L8: Repeated measurements

22 / 44

Basic Statistics for health researchers - L8: Repeated measurements 22 / 44

Estimating summary statistics "at once" (2/2)

```
library(data.table)
dtL.pp <- as.data.table(dL.pp)
dtL.pp[,mean(visual),by = "week"]
```

	week	V1
1:	0	54.84848
2:	4	52.45887

```
dtL.pp[,list(mu=mean(visual), sigma=sd(visual)),  
        by = c("treatment","week")]
```

treatment	week	mu	sigma
Active	0	54.42105	14.59718
Active	4	50.91228	15.81114
Placebo	0	55.26496	15.11872
Placebo	4	53.96581	15.90973

Basic Statistics for health researchers - L8: Repeated measurements

23 / 44

Testing the treatment effect: t-test (1/2)

Compute the change:

```
dW.ppP$visualDiff04 <- dW.ppP$visual4-dW.ppP$visual0
dW.ppA$visualDiff04 <- dW.ppA$visual4-dW.ppA$visual0
```

t-test:

```
t.test(x=dW.ppP$visualDiff04, y=dW.ppA$visualDiff04)
```

Welch Two Sample t-test

```
data: dW.ppP$visualDiff04 and dW.ppA$visualDiff04
t = 2.019, df = 223.08, p-value = 0.04469
alternative hypothesis: true difference in means is not equal to
95 percent confidence interval:
 0.05288768 4.36636558
sample estimates:
mean of x mean of y
-1.299145 -3.508772
```

Basic Statistics for health researchers - L8: Repeated measurements

24 / 44

Testing the treatment effect: t-test (2/2)

Underlying statistical model:

- $\text{visual}_{\text{week } 4, \text{Active}} - \text{visual}_{\text{week } 0, \text{Active}} \sim \mathcal{N}(\mu_{A,4-0}, \sigma_{A,4-0}^2)$
- $\text{visual}_{\text{week } 4, \text{Placebo}} - \text{visual}_{\text{week } 0, \text{Placebo}} \sim \mathcal{N}(\mu_{P,4-0}, \sigma_{P,4-0}^2)$

Note 1: strictly speaking, normality is not a requirement

Note 2:

- $\mu_{A,4-0} = \mu_{A,4} - \mu_{A,0}$
- $\sigma_{A,4-0}^2 = \sigma_{A,0}^2 + \sigma_{A,4}^2 - 2\rho_A \sigma_{A,4} \sigma_{A,0}$

Was it a good idea to work on the difference?

Testing the treatment effect: t-test (2/2)

Underlying statistical model:

- $\text{visual}_{\text{week } 4, \text{Active}} - \text{visual}_{\text{week } 0, \text{Active}} \sim \mathcal{N}(\mu_{A,4-0}, \sigma_{A,4-0}^2)$
- $\text{visual}_{\text{week } 4, \text{Placebo}} - \text{visual}_{\text{week } 0, \text{Placebo}} \sim \mathcal{N}(\mu_{P,4-0}, \sigma_{P,4-0}^2)$

Note 1: strictly speaking, normality is not a requirement

Note 2:

- $\mu_{A,4-0} = \mu_{A,4} - \mu_{A,0}$
- $\sigma_{A,4-0}^2 = \sigma_{A,0}^2 + \sigma_{A,4}^2 - 2\rho_A \sigma_{A,4} \sigma_{A,0}$

Basic Statistics for health researchers - L8: Repeated measurements

25 / 44

What about a standard linear regression?

```
dW.pp$visualDiff04 <- dW.pp$visual4 - dW.pp$visual0
lm(visualDiff04 ~ treatment, data = dW.pp)
```

Call:

```
lm(formula = visualDiff04 ~ treatment, data = dW.pp)
```

Coefficients:

(Intercept)	treatmentPlacebo
-3.509	2.210

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○○○
○○○○○○○○
○○○○○○○○

Multivariate model
○●○○○
○○○○○○○○
○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

Introduction
○○○○○
○○○

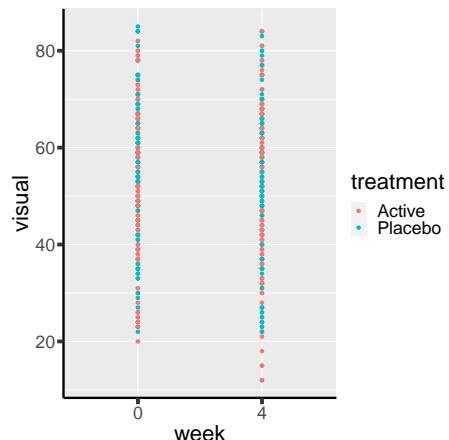
Pre-post study
○○○○○○○○○○
○○○○○○○○○○
○○○○○○○○○○

Multivariate model
○○●○○○
○○○○○○○○○○
○○○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

Graphical display: spaghetti plot (1/3)

```
gg <- ggplot(data = dL.pp,  
               mapping = aes(x=week, y=visual, color=treatment))  
gg + geom_point()
```



Basic Statistics for health researchers - L8: Repeated measurements

30 / 44

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○○○
○○○○○○○○
○○○○○○○○

Multivariate model
○○●○○○
○○○○○○○○
○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

Introduction
○○○○○
○○○

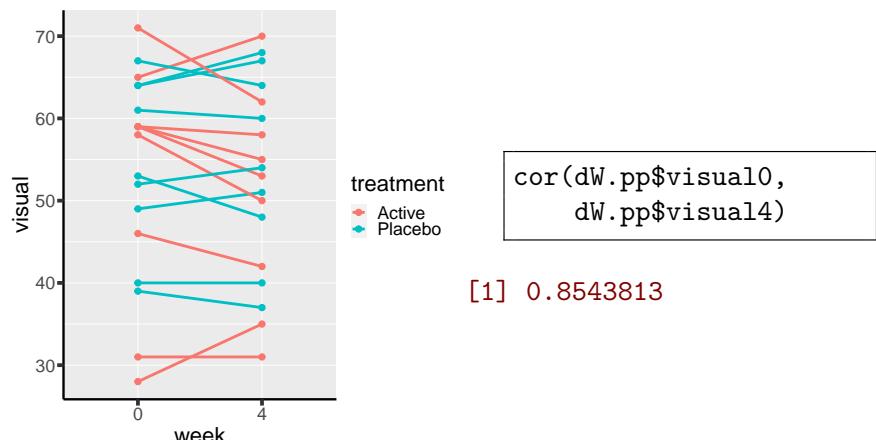
Pre-post study
○○○○○○○○○○
○○○○○○○○○○
○○○○○○○○○○

Multivariate model
○○●○○○
○○○○○○○○○○
○○○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

Graphical display: spaghetti plot (2/3)

```
gg <- ggplot(data = dL.pp[dL.pp$subject < 20,],  
               mapping = aes(x=week, y=visual, color=treatment))  
gg + geom_point() + geom_line(aes(group = subject))
```



cor(dW.pp\$visual0,
 dW.pp\$visual4)

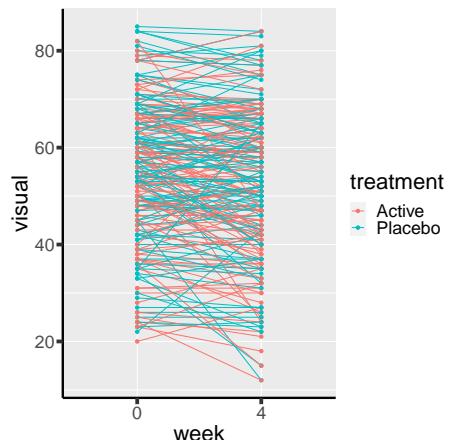
[1] 0.8543813

Basic Statistics for health researchers - L8: Repeated measurements

32 / 44

Graphical display: spaghetti plot (2/3)

```
gg <- ggplot(data = dL.pp,  
               mapping = aes(x=week, y=visual, color=treatment))  
gg + geom_point() + geom_line(aes(group = subject))
```



Basic Statistics for health researchers - L8: Repeated measurements

31 / 44

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○○○
○○○○○○○○
○○○○○○○○

Multivariate model
○○●○○○
○○○○○○○○
○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

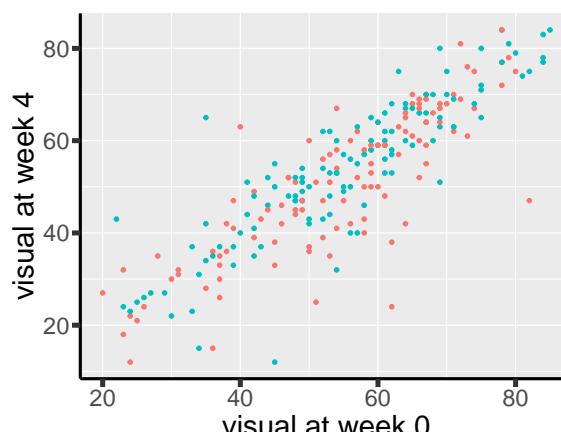
Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○○○
○○○○○○○○○○
○○○○○○○○○○

Multivariate model
○○●○○○
○○○○○○○○○○
○○○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

Bivariate distribution



Basic Statistics for health researchers - L8: Repeated measurements

treatment ● Active ■ Placebo

33 / 44

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○
○○○○○
○○○○

Multivariate model
○○○●
○○○
○○○○○

Conclusion
○○○
○○○○○○○○

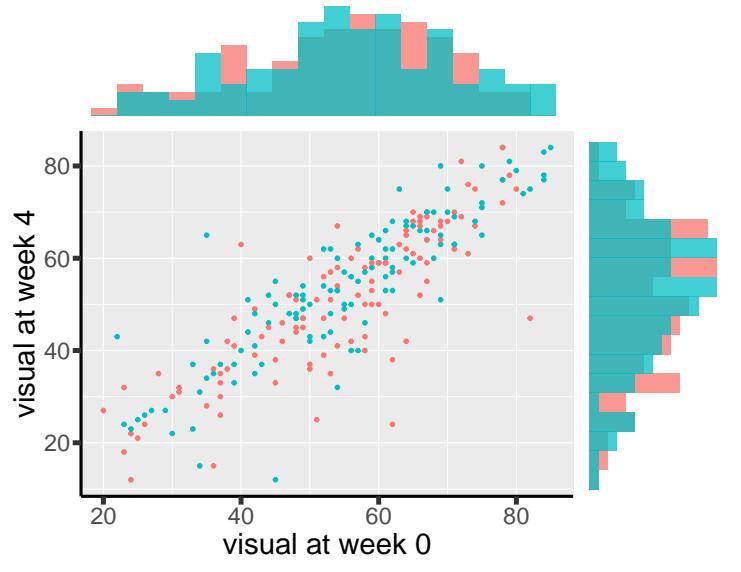
Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○
○○○○○
○○○○

Multivariate model
○○○●
○○○
○○○○○○○○

Conclusion
○○○
○○○○○○○○○○

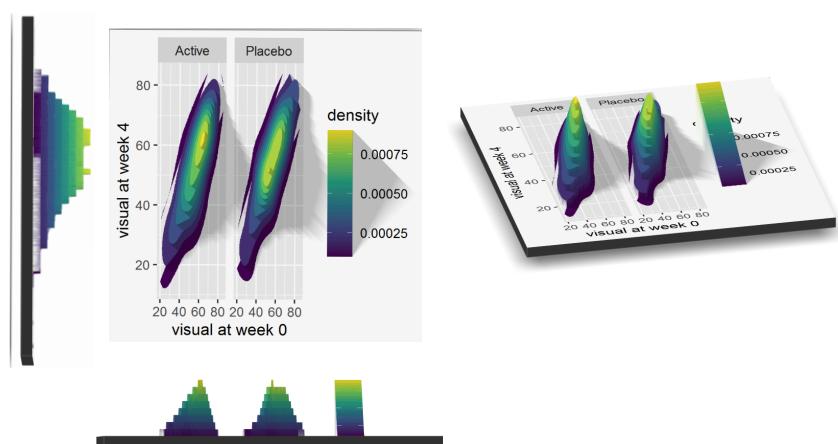
Bivariate distribution



Basic Statistics for health researchers - L8: Repeated measurements

33 / 44

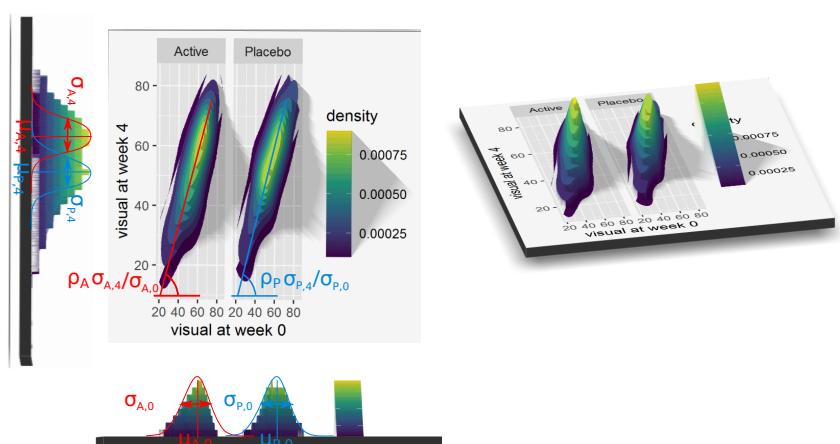
Bivariate distribution



Basic Statistics for health researchers - L8: Repeated measurements

33 / 44

Bivariate distribution



Basic Statistics for health researchers - L8: Repeated measurements

33 / 44

The general model (under normality assumption) in the pre-post study is:

- Active group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \dots \\ \dots \end{bmatrix}, \begin{bmatrix} \dots & \dots \\ \dots & \dots \end{bmatrix} \right)$$

- Placebo group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \dots \\ \dots \end{bmatrix}, \begin{bmatrix} \dots & \dots \\ \dots & \dots \end{bmatrix} \right)$$

34 / 44

Basic Statistics for health researchers - L8: Repeated measurements

Multivariate linear model - natural parametrisation

The general model (under normality assumption) in the pre-post study is:

- Active group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{A,0} \\ \mu_{A,4} \end{bmatrix}, \begin{bmatrix} \sigma_{A,0}^2 & \rho_A \sigma_{A,0} \sigma_{A,4} \\ \rho_A \sigma_{A,0} \sigma_{A,4} & \sigma_{A,4}^2 \end{bmatrix} \right)$$

- Placebo group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{P,0} \\ \mu_{P,4} \end{bmatrix}, \begin{bmatrix} \sigma_{P,0}^2 & \rho_P \sigma_{P,0} \sigma_{P,4} \\ \rho_P \sigma_{P,0} \sigma_{P,4} & \sigma_{P,4}^2 \end{bmatrix} \right)$$

Random intercept model - standard parametrisation

- Active group:

$$\begin{bmatrix} \mu_{A,0} = \text{visual}_{\text{week } 0} \\ \mu_{A,4} = \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{A,0} = \alpha \\ \mu_{A,4} = \alpha + \beta \end{bmatrix}, \begin{bmatrix} \sigma^2 + \tau & \tau \\ \tau & \sigma^2 + \tau \end{bmatrix} \right)$$

- Placebo group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{P,0} = \alpha + \gamma \\ \mu_{P,4} = \alpha + \beta + \gamma + \delta \end{bmatrix}, \begin{bmatrix} \sigma^2 + \tau & \tau \\ \tau & \sigma^2 + \tau \end{bmatrix} \right)$$

Can you re-express the null hypothesis using the model parameters?

Random intercept model - standard parametrisation

- Active group:

$$\begin{bmatrix} \mu_{A,0} = \text{visual}_{\text{week } 0} \\ \mu_{A,4} = \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{A,0} = \alpha \\ \mu_{A,4} = \alpha + \beta \end{bmatrix}, \begin{bmatrix} \sigma^2 + \tau & \tau \\ \tau & \sigma^2 + \tau \end{bmatrix} \right)$$

- Placebo group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{P,0} = \alpha + \gamma \\ \mu_{P,4} = \alpha + \beta + \gamma + \delta \end{bmatrix}, \begin{bmatrix} \sigma^2 + \tau & \tau \\ \tau & \sigma^2 + \tau \end{bmatrix} \right)$$

Can you re-express the null hypothesis using the model parameters?

$$\mathcal{H}_0: \delta = 0$$

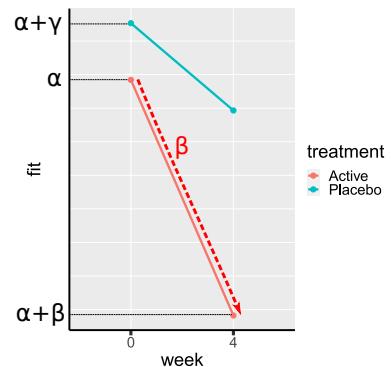
Random intercept model in R

```
e.lme <- lme(visual ~ week*treatment,
               random = ~1|subject,
               data = dL.pp)
logLik(e.lme)
```

'log Lik.' -1761.094 (df=6)

Display fitted values

```
gg <- ggplot(UX, aes(x = week, y = fit,
                      group = treatment, color = treatment))
gg + geom_point() + geom_line()
```

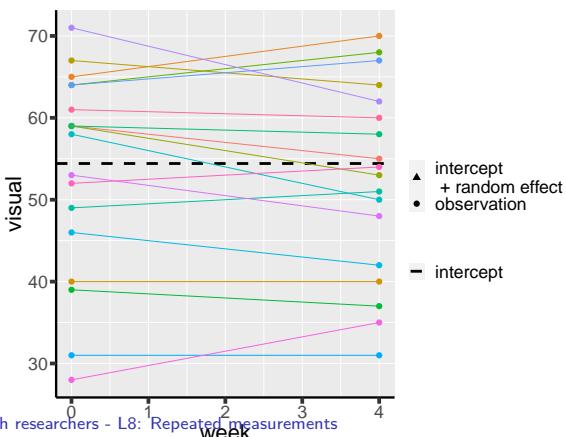


Variance-covariance structure (1/2)

For individual $i \in \{1, \dots, 231\}$ and week $t \in \{0, 4\}$:

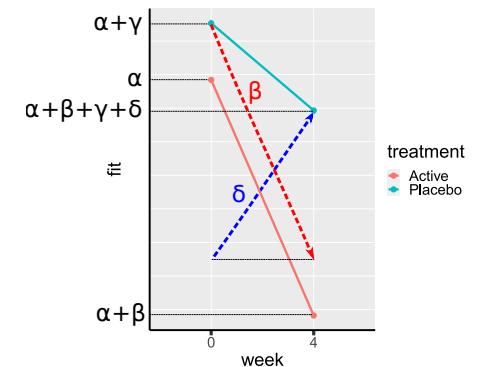
$$\varepsilon_{it} = u_i + \xi_{it}$$

where $u_i \sim \mathcal{N}(0, \tau)$ are independent random effects
 $\xi_{it} \sim \mathcal{N}(0, \sigma^2)$ are independent residuals



Display fitted values

```
gg <- ggplot(UX, aes(x = week, y = fit,
                      group = treatment, color = treatment))
gg + geom_point() + geom_line()
```

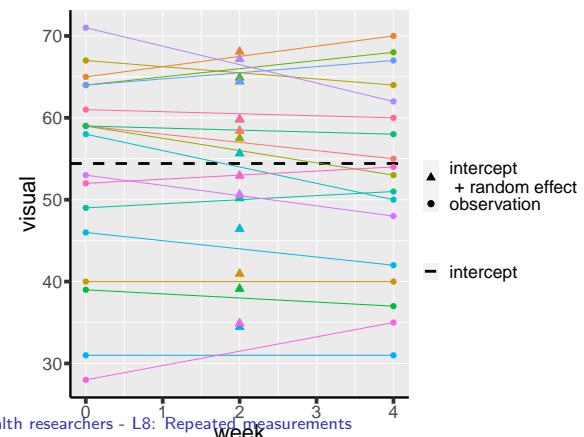


Variance-covariance structure (1/2)

For individual $i \in \{1, \dots, 231\}$ and week $t \in \{0, 4\}$:

$$\varepsilon_{it} = u_i + \xi_{it}$$

where $u_i \sim \mathcal{N}(0, \tau)$ are independent random effects
 $\xi_{it} \sim \mathcal{N}(0, \sigma^2)$ are independent residuals

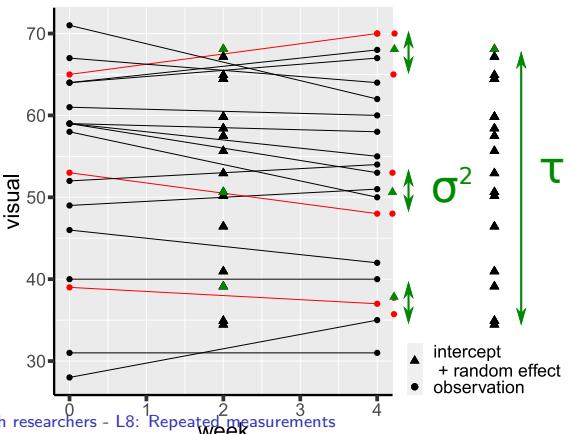


Variance-covariance structure (1/2)

For individual $i \in \{1, \dots, 231\}$ and week $t \in \{0, 4\}$:

$$\varepsilon_{it} = u_i + \xi_{it}$$

where $u_i \sim \mathcal{N}(0, \tau)$ are independent random effects
 $\xi_{it} \sim \mathcal{N}(0, \sigma^2)$ are independent residuals



Basic Statistics for health researchers - L8: Repeated measurements		Week	40 / 44
Introduction oo oo	Pre-post study oooooooo ooooo oooo	Multivariate model ooooo oo ooooo●	Conclusion ooo oooooooooooo

Variance-covariance structure (2/2)

```
getVarCov(e.lme, individuals = 2:3, type = "marginal")
```

```

subject 2
Marginal variance covariance matrix
      1         2
1 236.25 201.80
2 201.80 236.25
      Standard Deviations: 15.37 15.37
subject 3
Marginal variance covariance matrix
      1         2
1 236.25 201.80
2 201.80 236.25
      Standard Deviations: 15.37 15.37

```

What hypotheses are we making?

Basic Statistics for health researchers - L8: Repeated measurements

Variance-covariance structure (2/2)

```
getVarCov(e.lme, individuals = 2:3, type = "marginal")
```

```
subject 2
Marginal variance covariance matrix
      1         2
1 236.25 201.80
2 201.80 236.25
      Standard Deviations: 15.37 15.37
subject 3
Marginal variance covariance matrix
      1         2
1 236.25 201.80
2 201.80 236.25
      Standard Deviations: 15.37 15.37
```

Basic Statistics for health researchers - L8: Repeated measurements

41 / 44

What hypotheses are we making?

What hypotheses are we making?

$$\sigma_{A,0}^2 = \sigma_{A,4}^2 = \sigma_{P,0}^2 = \sigma_{P,4}^2 \quad \rho_A = \rho_P$$

Basic Statistics for Health researchers - L8: Repeated measurements

Basic Statistics for health researchers - L8: Repeated measurements

Basic Statistics for health researchers - L8: Repeated measurements 41 / 44

41 / 44

Conclusion

Collecting several measurements per subject is a good idea:

- reduce uncertainty/confounding (each subject is its own control)
- give more insight into the treatment effect
- scheduled measurement time is recommended

But is also challenging:

- more demanding for the patient ( drop-out!)
- more complex to organize
(e.g. ensure subjects follow the schedule)
- often require dedicated/advanced statistical tools

42 / 44

Introduction	Pre-post study	Multivariate model	Conclusion
○○○○ ○○	○○○○○○ ○○○○ ○○○	○○○ ○○ ○○○○○	○○● ○○○○○○○○

Want to know more?

Ph.D. course:

- Statistical analysis of correlated and repeated measurements
(course director: Julie Forman)

Contents

This course is concerned with analysis of correlated quantitative data arising e.g. when taking observations from clusters of subjects, repeatedly over time on the same subjects, or by applying different treatment to different parts of the body. Pitfalls of traditional statistical analyses will be discussed and appropriate models for the analysis of e.g. baseline follow-up studies, cross-over studies, and cluster randomized trials will be exemplified.

For supplementary reading we recommend:

- FLW: G.M. Fitzmaurice, N.M. Laird and J.H. Ware, Applied Longitudinal Analysis (2nd edition), John Wiley & sons, 2011.

Please note that the book is available as e-book on KB (free download for KU students).

Day	Topics	Suggested reading*
1	Introduction to repeated measurements and clustered data. Basic theory of linear mixed models. Analysis of single group studies. Handling repeated measurements in SAS/R.	FLW 1.3. Tutorial 1.
2	Longitudinal data analysis. Models for balanced and unbalanced designs. Analysis of randomized baseline follow-up studies.	FLW 5.7. Tutorial 2.
3	Analysis of clustered data. Variance components. Multi-level models. The linear growth model.	FLW 8, 21 & 22.
4	Select topics in linear mixed models. Cross-over studies. Repeatability and reproducibility of measurement methods.	Lecture notes only.
5	Models for binary and count data. Generalized linear mixed models. Marginal models and generalized estimating equations.	FLW 10-16
6	Missing data. Consequences and statistical handling.	FLW 17-18

44 / 44

Pro:

- generalize to more than two timepoints
- applicable in continuous time
- efficiently handles missing values
- more interpretable
(e.g. impose constant or linear treatment effect over time)

Cons:

- make more assumptions
- more difficult to describe (in an article)
- more complex to estimate and to perform statistical inference
(p-value, confidence intervals can be less reliable in small samples)

43 / 44

Introduction	Pre-post study	Multivariate model	Conclusion
○○○○ ○○	○○○○○○ ○○○○ ○○○	○○○ ○○ ○○○○○	○○● ○○○○○○○○

43 / 44

Introduction	Pre-post study	Multivariate model	Conclusion
○○○○ ○○	○○○○○○ ○○○○ ○○○	○○○ ○○ ○○○○○	○○● ○○○○○○○○

Reference I

(1997). Interferon Alfa-2a Is Ineffective for Patients With Choroidal Neovascularization Secondary to Age-Related Macular Degeneration: Results of a Prospective Randomized Placebo-Controlled Clinical Trial. *Archives of Ophthalmology*, 115(7):865–872.

Beliveau, V., Ganz, M., Feng, L., Ozenne, B., Højgaard, L., Fisher, P. M., Svarer, C., Greve, D. N., and Knudsen, G. M. (2017). A high-resolution *in vivo* atlas of the human brain's serotonin system. *Journal of Neuroscience*, 37(1):120–128.

Hjordt, L. V., Ozenne, B., Armand, S., Dam, V. H., Jensen, C. G., Köhler-Forsberg, K., Knudsen, G. M., and Stenbæk, D. S. (2020). Psychometric properties of the verbal affective memory test-26 and evaluation of affective biases in major depressive disorder. *Frontiers in psychology*, 11:961.

45 / 44

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○

Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
●○○○○○○○○

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○○

Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
○●○○○○○○○

Reference II

Van Reeth, E., Tham, I. W., Tan, C. H., and Poh, C. L. (2012).
Super-resolution in magnetic resonance imaging: a review.
Concepts in Magnetic Resonance Part A, 40(6):306–325.

Natural parametrisation (1/3)

Shorten names:

```
dL.pp$treat <- factor(dL.pp$treatment,  
levels = c("Active", "Placebo"),  
labels = c("A", "P"))
```

```
e2.lme <- lme(visual ~ week:treat-1,  
random = ~1|subject,  
data = dL.pp)  
logLik(e2.lme)
```

'log Lik.' -1761.094 (df=6)

```
fixef(e2.lme)
```

week0:treatA	week4:treatA	week0:treatP	week4:treatP
54.42105	50.91228	55.26496	53.96581

Basic Statistics for health researchers - L8: Repeated measurements

Basic Statistics for health researchers - L8: Repeated measurements

46 / 44

47 / 44

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○

Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
●○○○○○○○○

Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○
○○○○
○○○○○

Multivariate model
○○○○
○○
○○○○○

Conclusion
○○○
○●○○○○○○○

Natural parametrisation (2/3)

```
library(multcomp)  
C <- "week4:treatA-week0:treatA-week4:treatP+week0:  
      treatP=0"  
e2.glht <- glht(e2.lme, linfct = C)  
rownames(e2.glht$linfct) <- "delta"  
summary(e2.glht)
```

Simultaneous Tests for General Linear Hypotheses

```
Fit: lme.formula(fixed = visual ~ week:treat - 1, data = dL.pp,  
                  subject)
```

Linear Hypotheses:

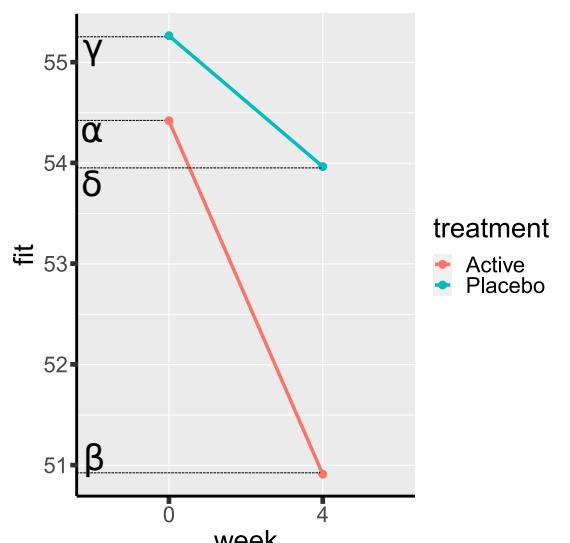
Estimate	Std. Error	z value	Pr(> z)
delta == 0	-2.210	1.092	-2.023 0.0431 *

Signif. codes:	0 *** 0.001 ** 0.01 * 0.05 . 0.1 ' ' 1		
(Adjusted p values reported -- single-step method)			

Basic Statistics for health researchers - L8: Repeated measurements

48 / 44

Natural parametrisation (3/3)



Basic Statistics for health researchers - L8: Repeated measurements

49 / 44



Testing multiple null hypotheses

```

library(multcomp)
e.glht <- glht(e.lme,
linfct = c("week4:treatmentPlacebo=0",
          "treatmentPlacebo+week4:treatmentPlacebo=0"))
rownames(e.glht$linfct) <- c("change","final value")
summary(e.glht)

```

Simultaneous Tests for General Linear Hypotheses

```
Fit: lme.formula(fixed = visual ~ week * treatment, data = dL.pp  
      random = ~1 | subject)
```

Linear Hypotheses:

```

Estimate Std. Error z value Pr(>|z|)
change == 0      2.210      1.092   2.023   0.0828 .
final value == 0 3.054      2.023   1.510   0.2397
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

```

*² tests for health researchers - I: Repeated measurements

Basic Statistics for health researchers - 18: Repeated measurements



Multivariate linear model in R (no missing value)

Parametrisation

- Active group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{A,0} \\ \mu_{A,4} \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho \sigma^2 k_{A,4} \\ \rho \sigma^2 k_{A,4} & \sigma^2 k_{A,4}^2 \end{bmatrix} \right)$$

- Placebo group:

$$\begin{bmatrix} \text{visual}_{\text{week } 0} \\ \text{visual}_{\text{week } 4} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_{P,0} \\ \mu_{P,4} \end{bmatrix}, \begin{bmatrix} \sigma^2 k_{P,0} & \rho \sigma^2 k_{P,0} k_{P,4} \\ \rho \sigma^2 k_{P,0} k_{P,4} & \sigma^2 k_{P,4}^2 \end{bmatrix} \right)$$



```
e.gls <- gls(visual ~ week*treatment, data = dL.pp,
              correlation = corSymm(form=~1|subject),
              weight = varIdent(form=~1|treatment*week))
```

Correlation parameter (ρ):

```
coef(e.gls$modelStruct$corStruct, unconstrained=FALSE)
```

[1] 0.8571637

Reference variance parameter (σ^2)

Standard deviation inflation factor ($\{k_{A,4}, k_{P,0}, k_{P,4}\}$):

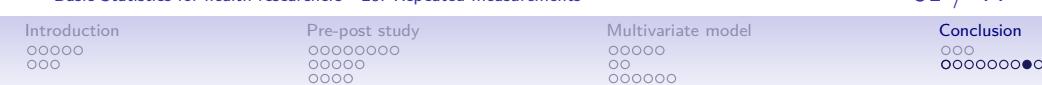
```
sigma(e.lme)^2  
coef(e.gls$modelStruct$varStruct, unconstrained=FALSE)
```

[1] 34.45639

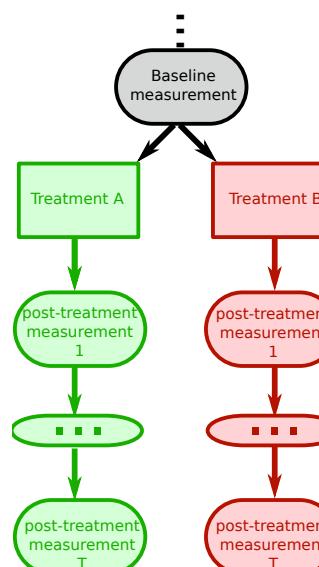
Active*4 Placebo*0 Placebo*4

1.0831625 0.9644284 1.0149006 Basic Statistics for Health Researchers I: Descriptive Measurements

Basic Statistics for health researchers - L8: Repeated measurements



Design: longitudinal study



Introduction
○○○○○
○○○

Pre-post study
○○○○○○○○
○○○○○
○○○○

Multivariate model
○○○○○
○○
○○○○○

Conclusion
○○○
○○○○○○○●

Design: cross-over study

