Basic Statistic for health researchers
Lecture 8: repeated measurements

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

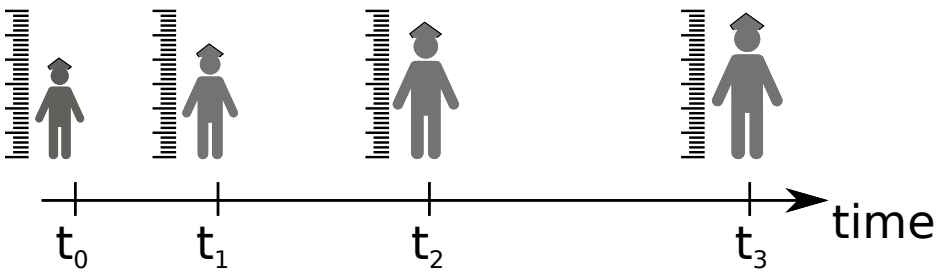
[1] Section of Biostatistics, Department of Public Health, University of Copenhagen
[2] Neurobiology Research Unit, University Hospital of Copenhagen, Rigshospitalet.

22-11-2021

Longitudinal study

Study design where one (or several) variable(s) are measured at different occasions on an experimental unit:

- typically: **outcome** measured on the **same patient** at **different timepoints**.

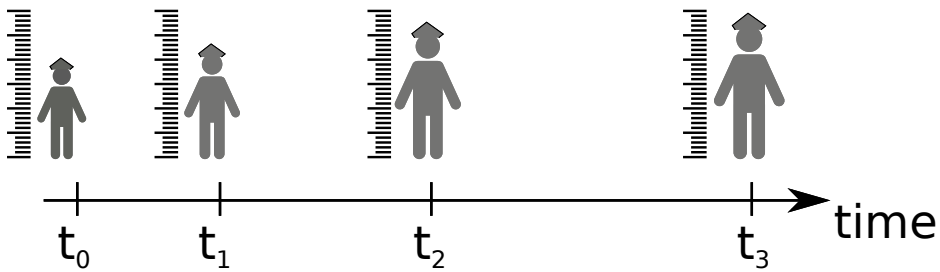


$\rightarrow$ assess time trends
or treatment effects (when there is a control group)

Longitudinal study

Study design where one (or several) variable(s) are measured at different occasions on an experimental unit:

- typically: **outcome** measured on the **same patient** at **different timepoints**.

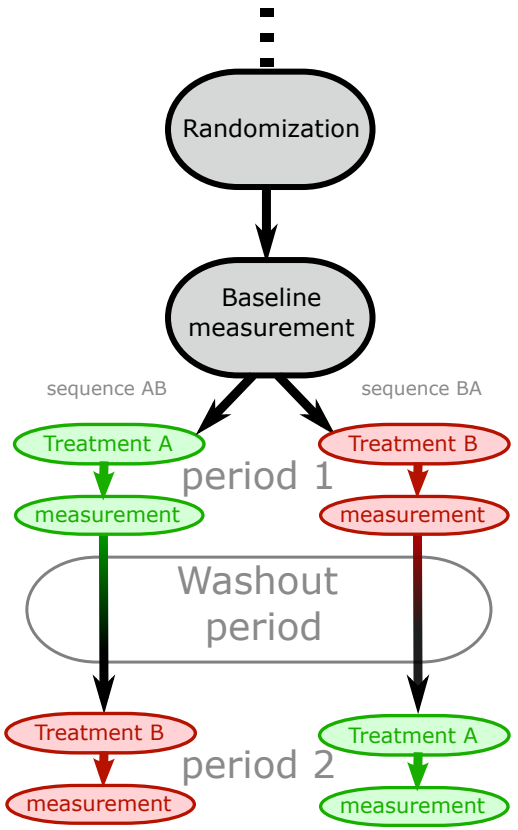


$\rightarrow$ assess time trends
or treatment effects (when there is a control group)

Timepoints are fixed by design
⚠ otherwise much more complex analysis

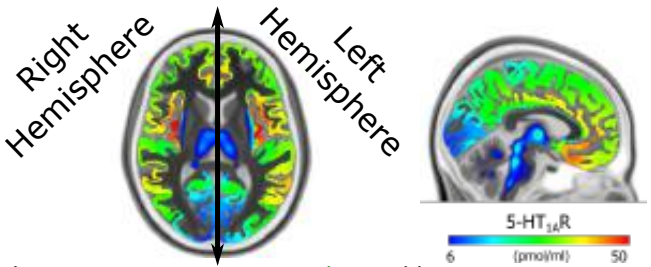
Other designs involving repeated measurements (1/2)

- cross-over: **outcome** measured on the **same patient** under **different treatments**.



Other designs involving repeated measurements (2/2)

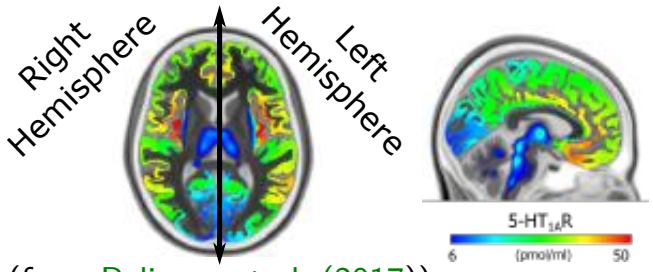
- the **same** type of **measurement** on the **same patient** at **different locations**.



(from Beliveau et al. (2017))

Other designs involving repeated measurements (2/2)

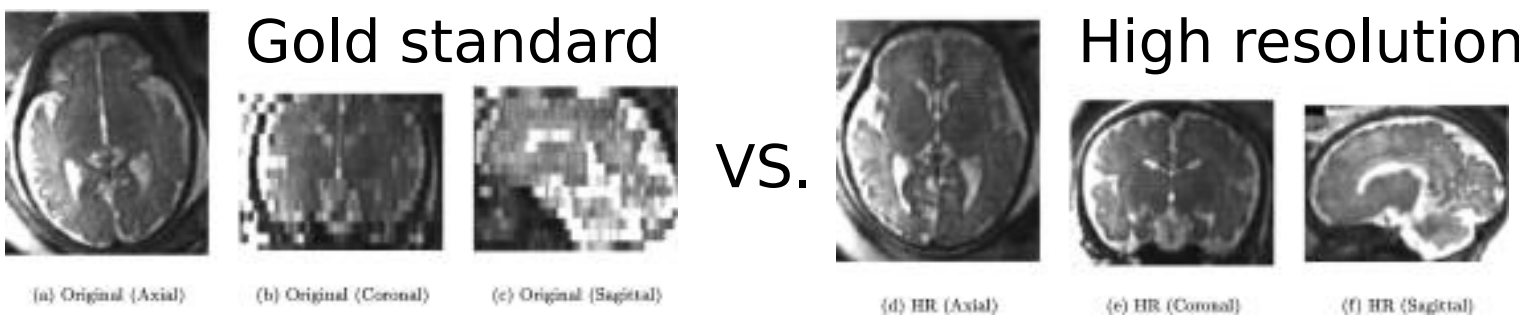
- the **same** type of **measurement** on the **same patient** at **different locations**.



(from Beliveau et al. (2017))

- test re-test study: **different** ways of **measuring** the same quantity on the **same patient**.

$\rightarrow$ assess the stability of a measurement device

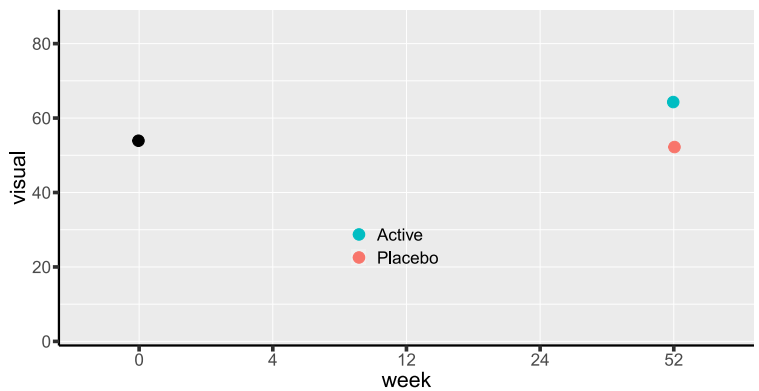


(from Van Reeth et al. (2012))

Why using repeated measurements? (1/4)

Why using repeated measurements? (1/4)

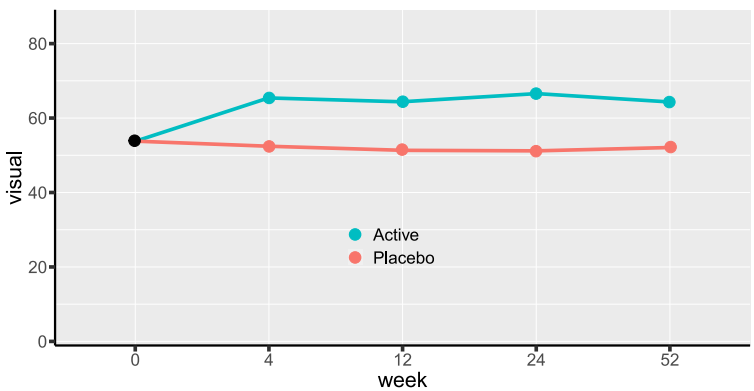
To **better understand** the time-dynamic of the **treatment effect**:



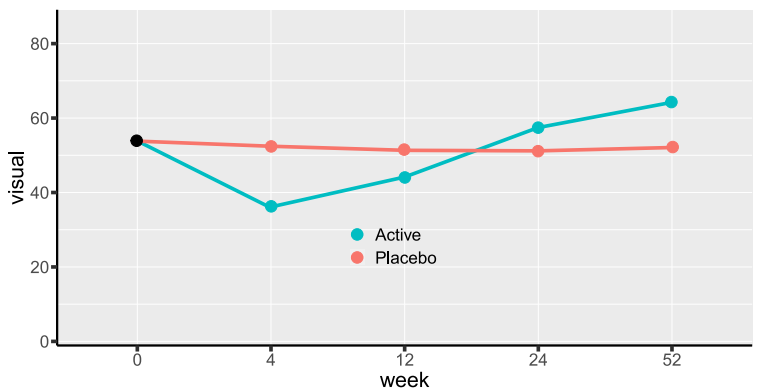
Why using repeated measurements? (1/4)

To **better understand** the time-dynamic of the **treatment effect**:

Constant effect



Time varying effect

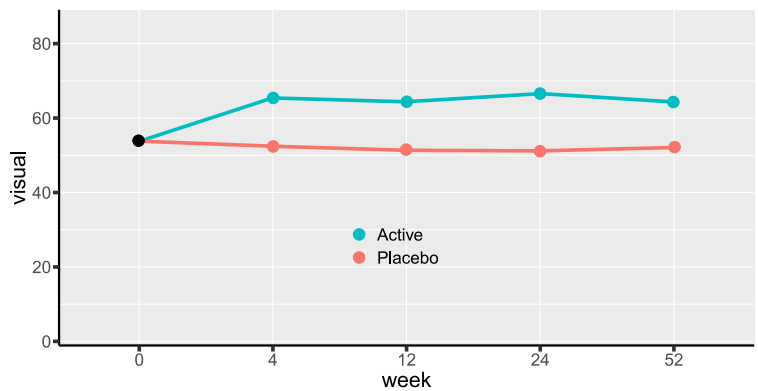


Why using repeated measurements? (1/4)

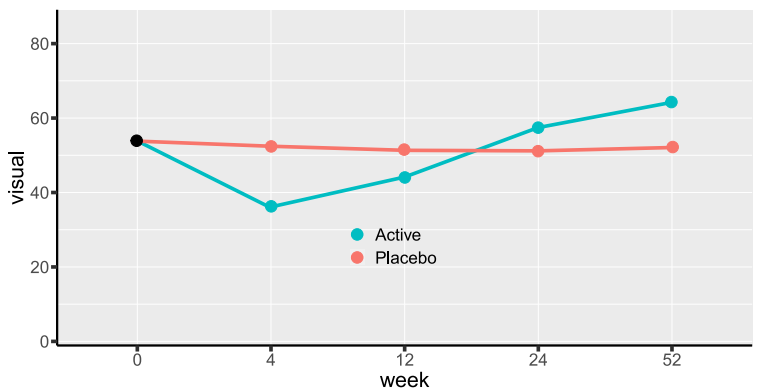
To **better understand** the time-dynamic of the **treatment effect**:

- is there any treatment effect?
- is there a sustained treatment effect?
- is there an immediate treatment effect?
- how do side effects occur after treatment intake ?

Constant effect



Time varying effect



Why using repeated measurements? (2/4)

To **save resources**:

- A baseline+post-treatment study will typically require [1] fewer subjects compared to a post-treatment only study.
  ...but requires 2 measurements instead of 1 per subject!

- A single study can be used for answering several research questions
  e.g., Immediate treatment effect on the serotonin system.
  Long-term treatment effect on depression.

[1] to achieve the same statistical power

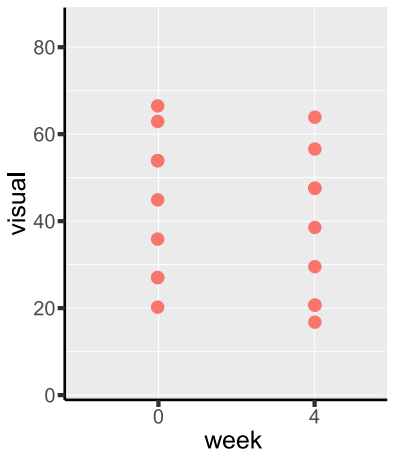
Why using repeated measurements? (3/4)

To **get a better estimation** of the treatment effect, especially in observational studies:

- idea: "use each patient as its own control"

$\rightarrow$ account for some confounders:
less bias

$\rightarrow$ account for some risk factors:
less uncertainty



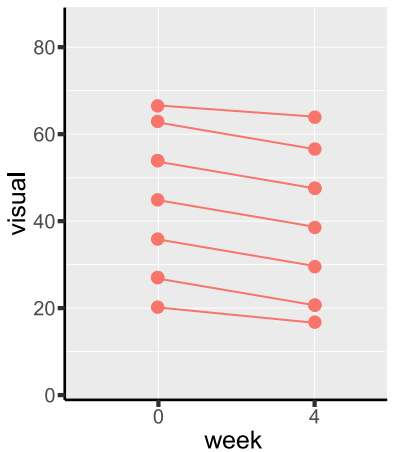
Why using repeated measurements? (3/4)

To **get a better estimation** of the treatment effect, especially in observational studies:

- idea: "use each patient as its own control"

$\rightarrow$ account for some confounders:
less bias

$\rightarrow$ account for some risk factors:
less uncertainty



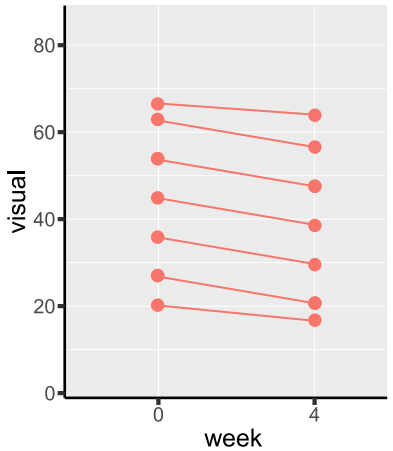
Why using repeated measurements? (3/4)

To **get a better estimation** of the treatment effect, especially in observational studies:

- idea: "use each patient as its own control"

$\rightarrow$ account for some confounders:
less bias

$\rightarrow$ account for some risk factors:
less uncertainty



⚠ Confounders/risk factors changing across repetitions:

- type of device used to make the measurement
- external events, e.g. food intake when monitoring glucose level

require special attention.

Why using repeated measurements? (4/4)

To **better handle missing values**:

- as the follow-up time increases, patient are more likely to drop-out
- regular follow-up can help:
  - to understand the reason(s) for drop-out
  - to limit the loss in statistical power due to drop-out
  - to adjust the analysis for informative drop-out

Challenges when analyzing repeated measurements (1/2)

Challenges when analyzing repeated measurements (1/2)

The usual **assumption of independent observations** is **violated**.
If ignored, this can lead to:

- ✘ incorrect p-values/confidence intervals (almost always)
- ✘ biased estimates (unless certain assumptions are met)

Challenges when analyzing repeated measurements (1/2)

The usual **assumption of independent observations** is **violated**.
If ignored, this can lead to:

- ✘ incorrect p-values/confidence intervals (almost always)
- ✘ biased estimates (unless certain assumptions are met)

Possible solutions:

- **summary-statistic**: summarize repetitions into one number (e.g. average, area under the curve, peak value)
- **univariate**: perform separate analyses at each timepoint.
- **multivariate**: simultaneously analyze all timepoints

Challenges when analyzing repeated measurements (2/2)

Multiple testing:

- several ways to test the treatment effect
  at any time, short term, long term, ...
- several possible statistical approaches
  multiple t-tests, random intercept model with linear / non-linear treatment effect ...

Challenges when analyzing repeated measurements (2/2)

Multiple testing:

- several ways to test the treatment effect
  at any time, short term, long term, ...
- several possible statistical approaches
  multiple t-tests, random intercept model with linear / non-linear treatment effect ...

Handling missing values:

- Drop-out (patients leaving the study)
  reason for leaving $\rightarrow$ how it should be handled
- Competing risks, e.g. patient died during the study
  informative censoring: complete case analysis usually wrong

Outline

✔ Introduction to repeated measurements
- definition and examples of study design
- challenges for the statistical analysis

Example of longitudinal study

Univariate approach

Multivariate approach

Illustrative example

- descriptive statistics and plots for repeated measurements
- concerns due to the presence of missing values
- what is a long and wide format

Illustration: ARMD trial (int, 1997)

Age-Related Macular Degeneration (ARMD) Trial:

- comparing interferon-α and placebo
- outcome $Y(t)$: change in vision over time

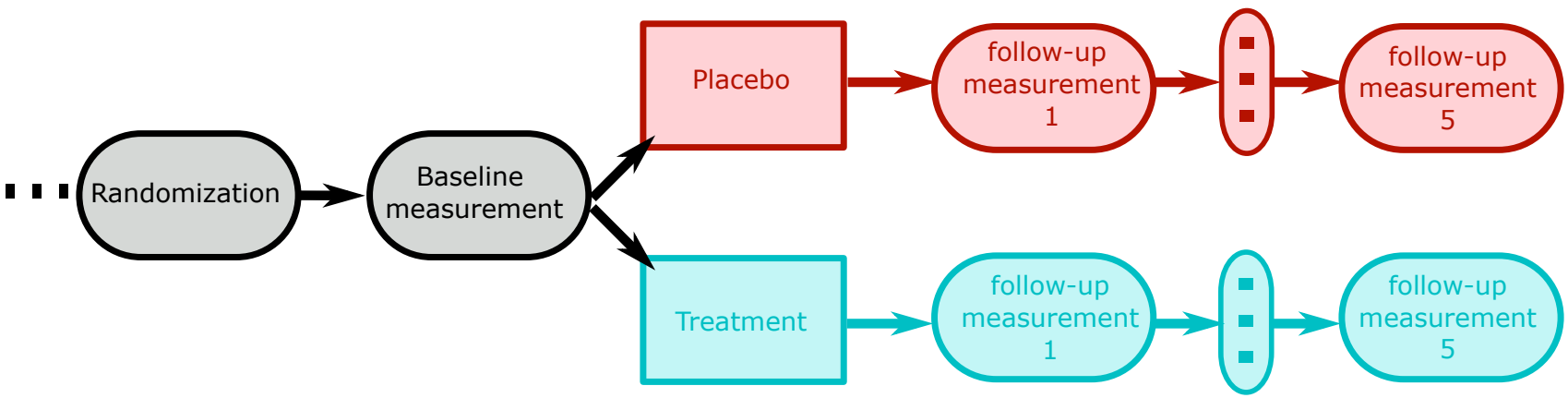
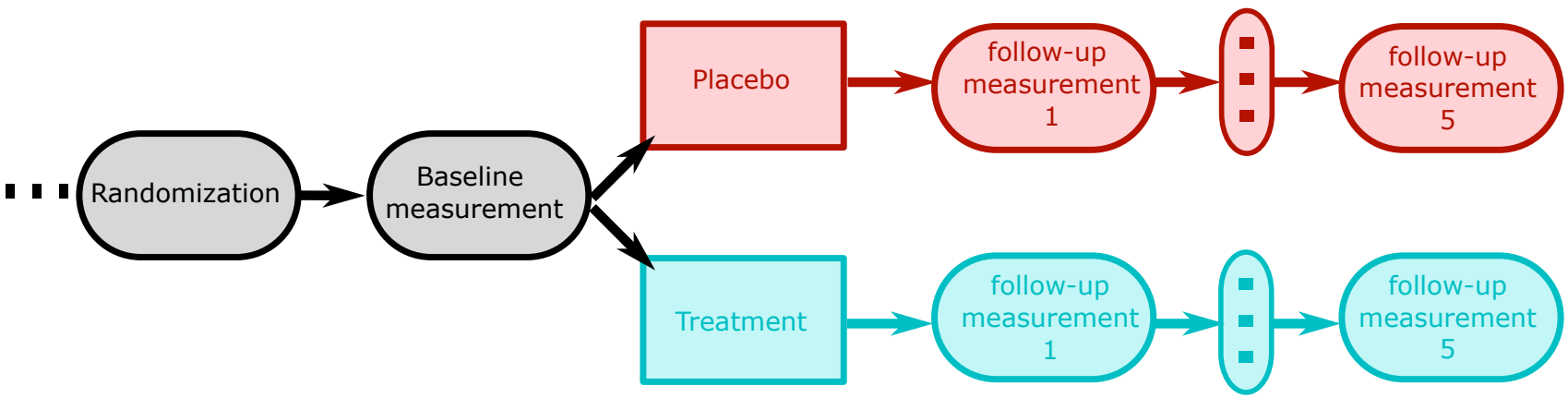


Illustration: ARMD trial (int, 1997)

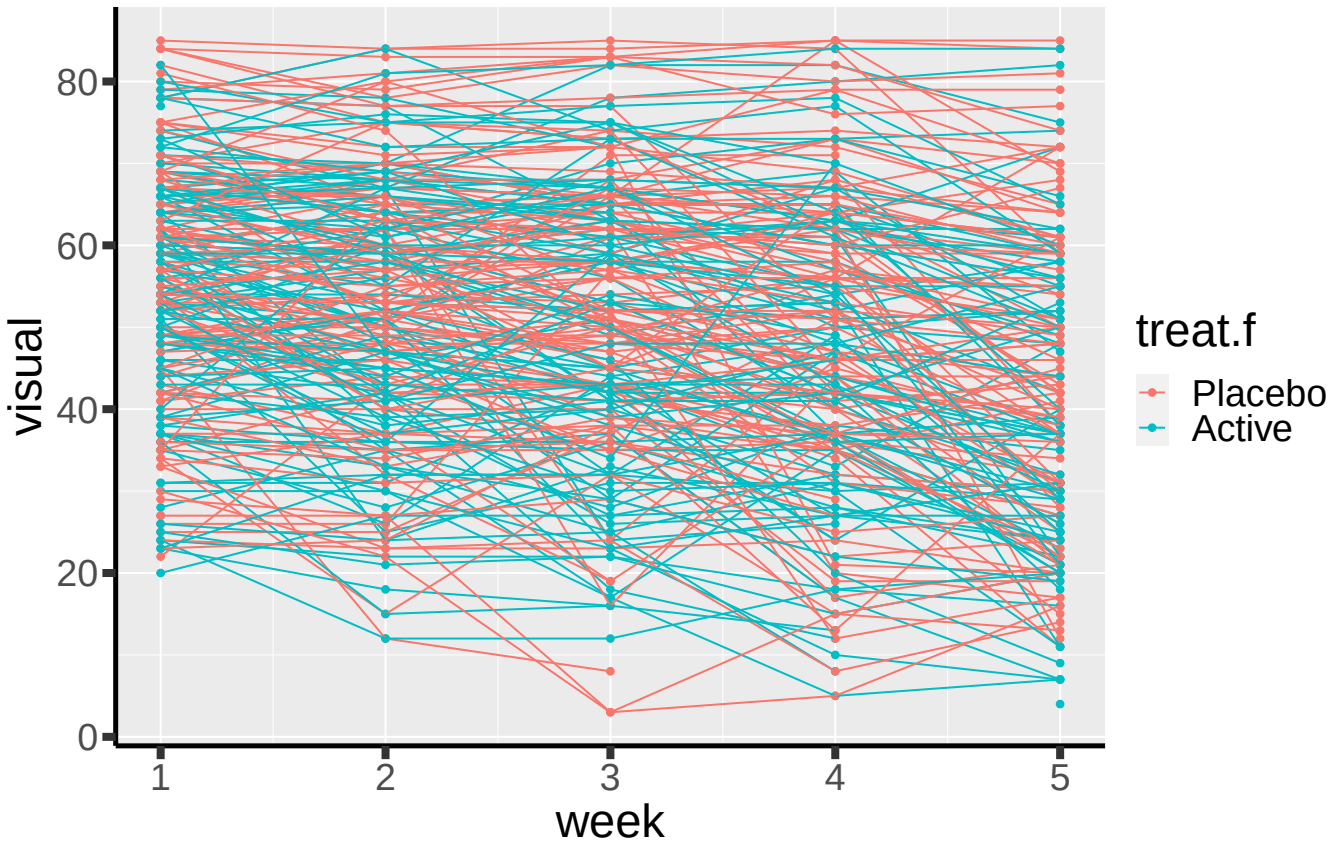
Age-Related Macular Degeneration (ARMD) Trial:

- comparing interferon-α and placebo
- outcome $Y(t)$: change in vision over time

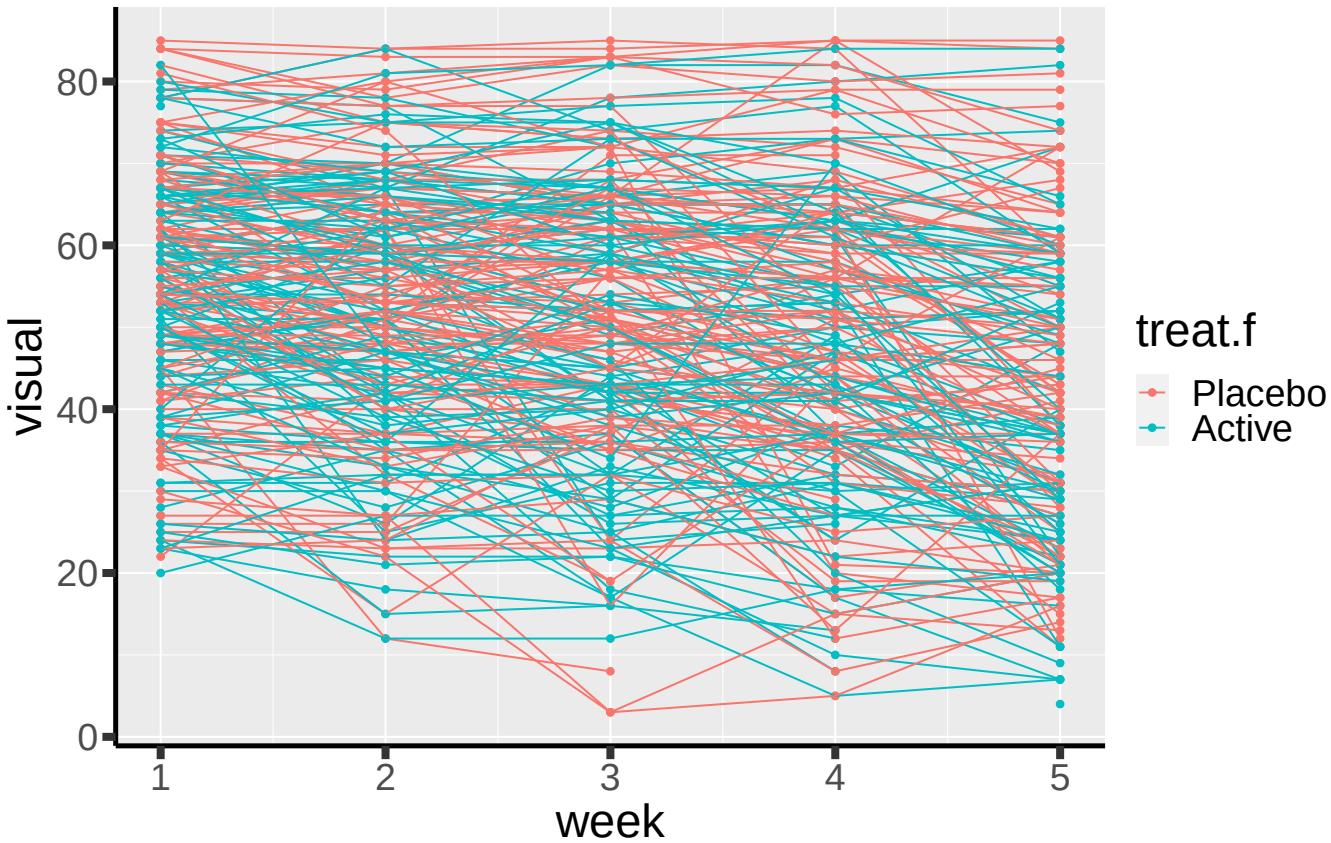


- **cluster** variable: subject (5 observations per cluster)
  $\rightarrow$ independent outcome replicates at the cluster level
- **repetition** variable: time

Visualizing the data: spaghetti plot



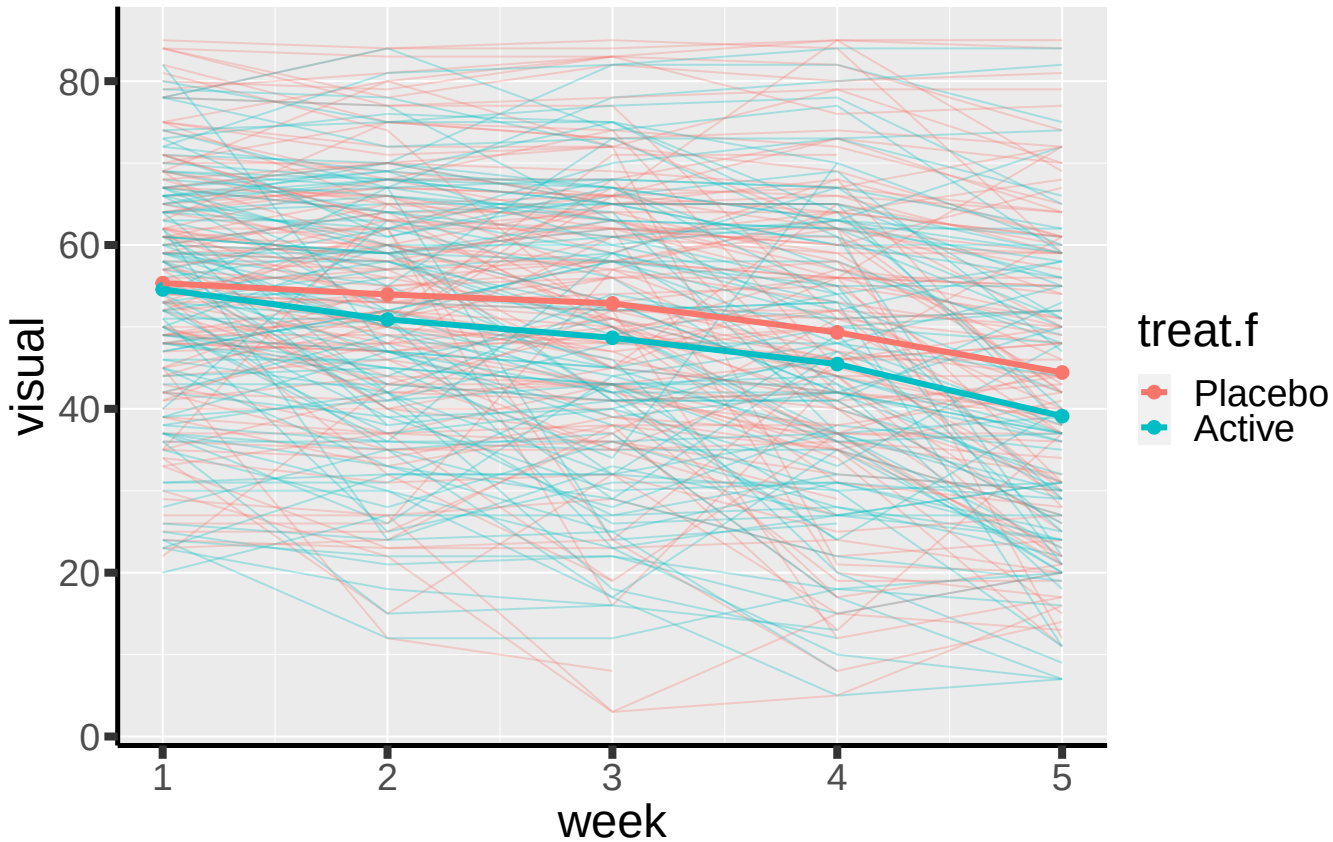
Visualizing the data: spaghetti plot



How can we summarize/describe the data?

Summary statistics (1/3)

- using the mean by group and timepoint:

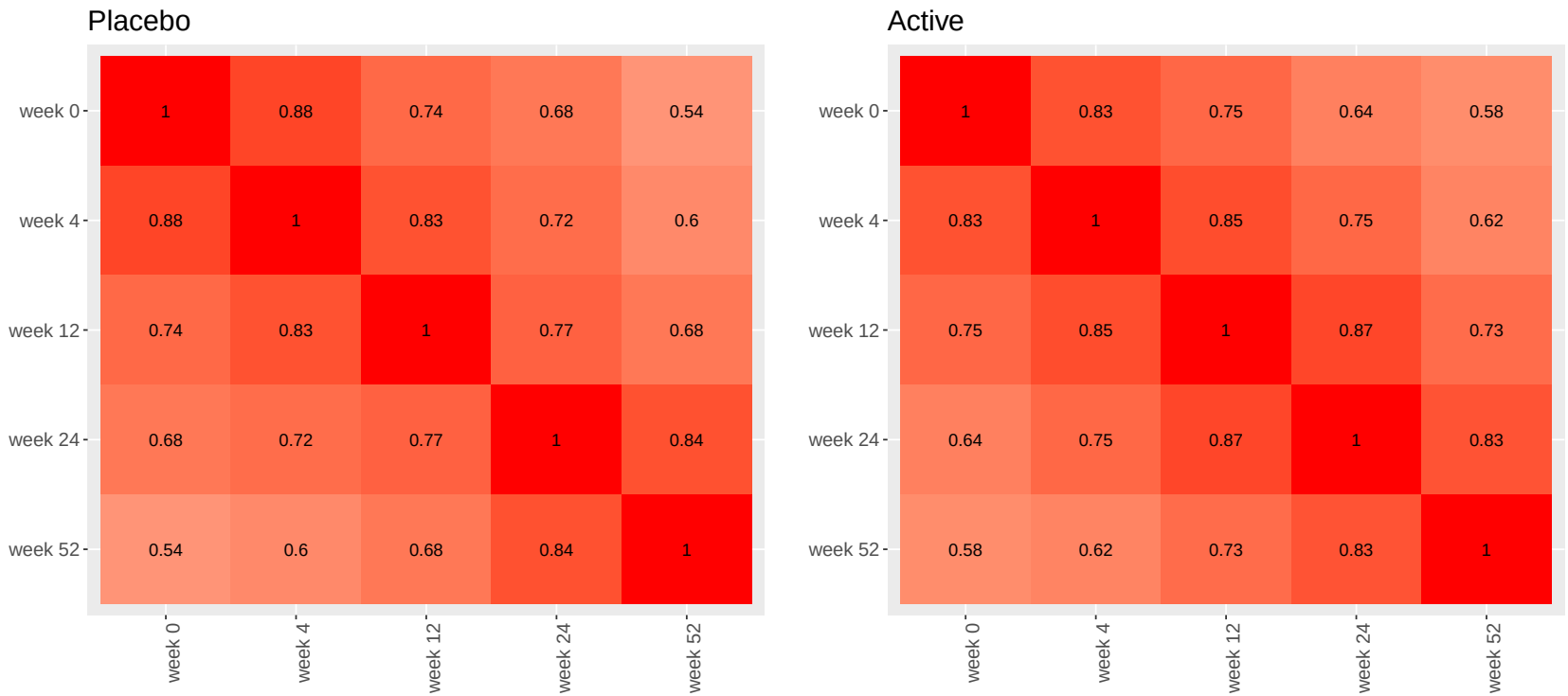


Summary statistics (2/3)

- dispertion over time (standard deviation)

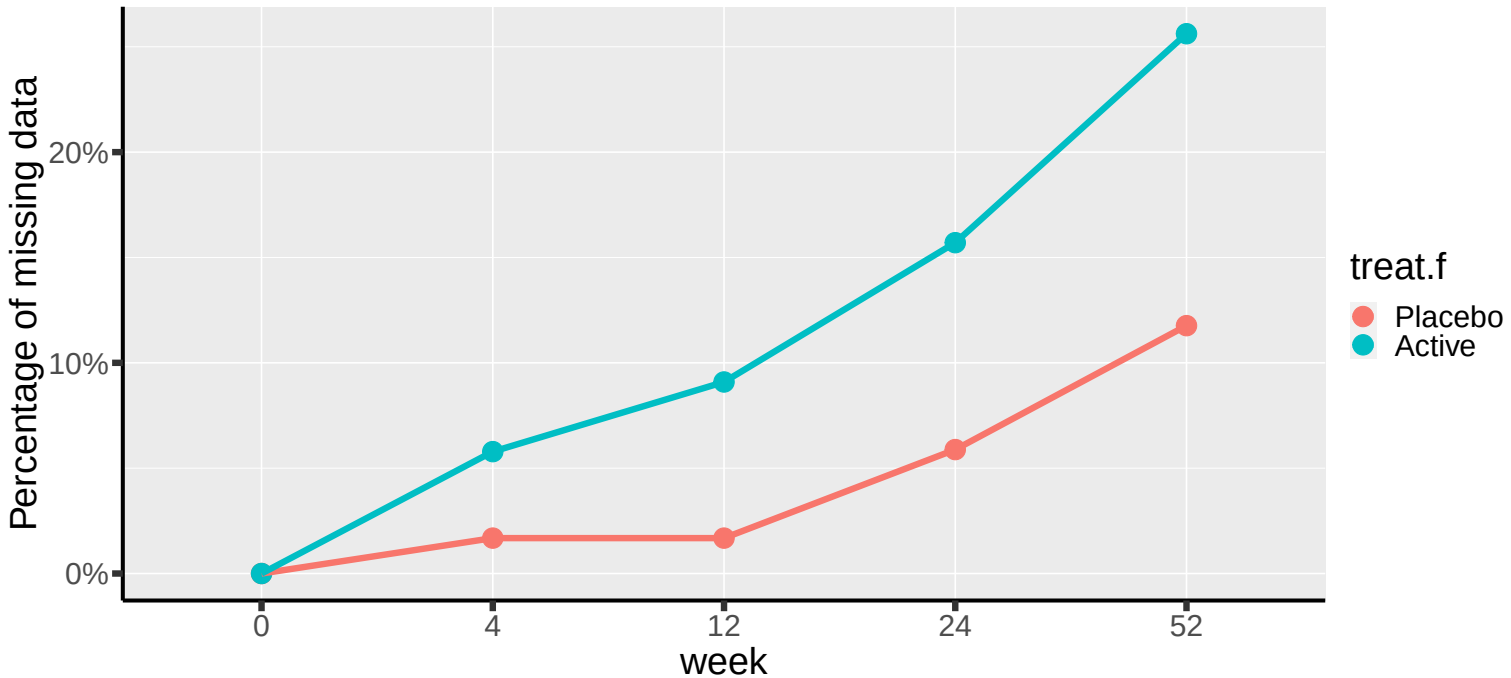
	visual0	visual4	visual12	visual24	visual52
Placebo	15.33143	15.38915	16.51203	18.61137	18.68844
Active	14.32523	15.99285	17.35207	17.84161	18.36214

- dependency in visual acuity over time (Pearson correlation)



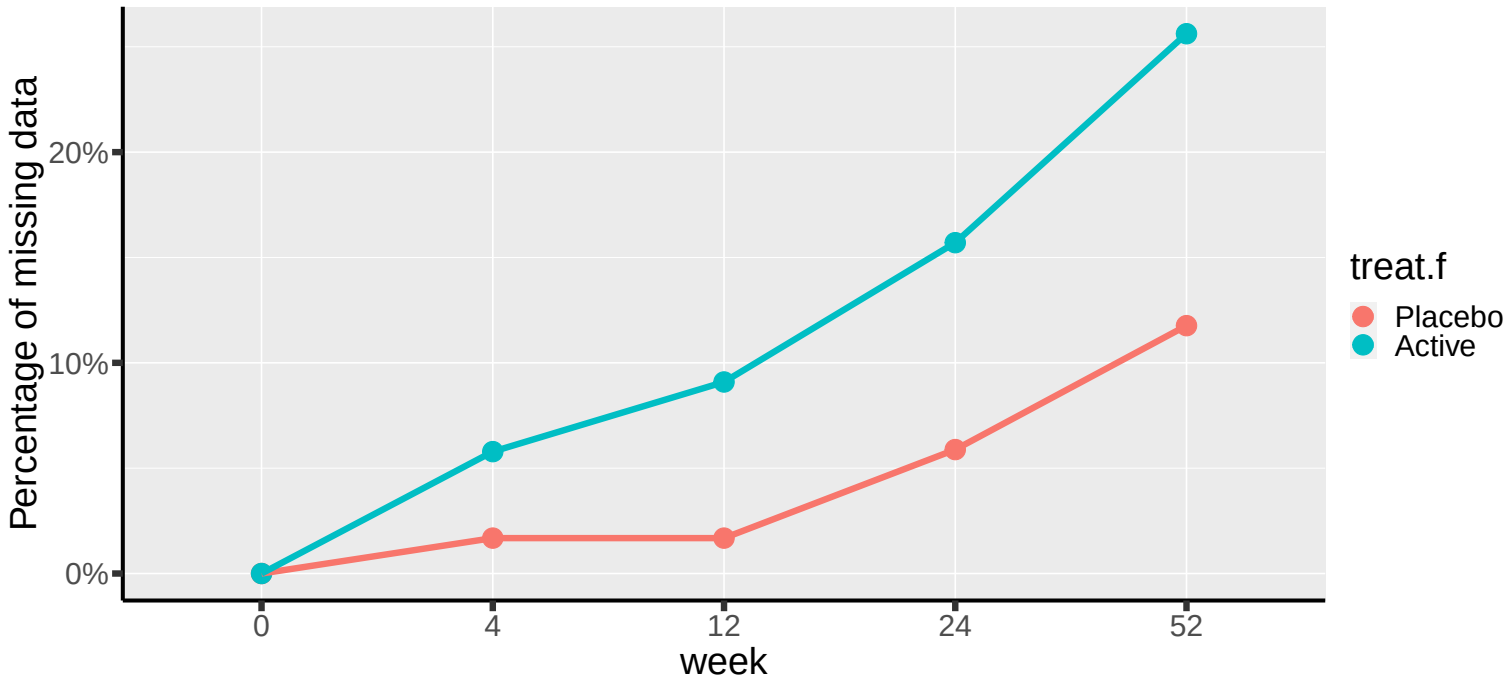
Summary statistics (3/3)

- what about missing values?



Summary statistics (3/3)

- what about missing values?



Concerns:

- treatment side effect(s) not measured by the outcome
- missing not at random may bias the estimated mean
  (upward bias if patients with weak vision are more likely to drop)

Wide vs. long format

Data in the wide format:

- 1 row = 1 subject ("level 1 data")

$\rightarrow$ independent replicate of $(Y(0), Y(4), Y(12), Y(24), Y(52))$

- convenient when working with one or two timepoints

	subject	treat.f	visual0	visual4	"..."	visual52
1	1	Active	59	55	...	NA
2	2	Active	65	70	...	55
3	3	Placebo	40	40	...	NA
4	4	Placebo	67	64	...	68
5	5	Active	70	NA	...	NA
6	6	Active	59	53	...	42

Wide vs. long format

Data in the long format:

- 1 row = 1 measurement of 1 subject ("level 0 data")
- convenient when performing operations over all timepoints

	subject	treat.f	week	visual
1	1	Active	0	59
2	1	Active	4	55
3	1	Active	12	45
4	1	Active	24	NA
5	1	Active	52	NA
6	2	Active	0	65
7	2	Active	4	70
8	2	Active	12	65
9	2	Active	24	65
10	2	Active	52	55

Technical note: summary statistic in the long format

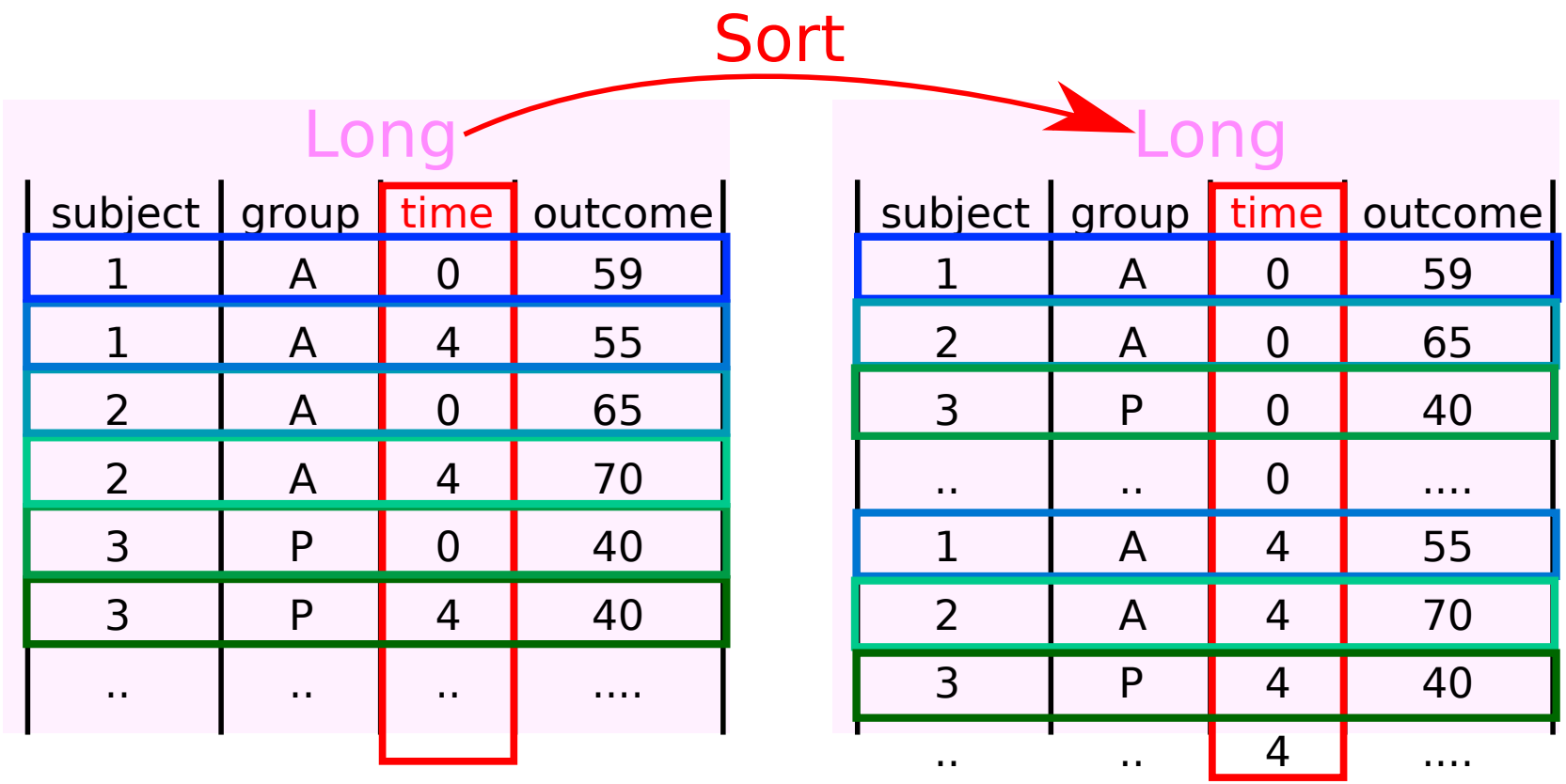
Example: average outcome value **at each timepoint**

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
..	..	..	....

Technical note: summary statistic in the long format

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

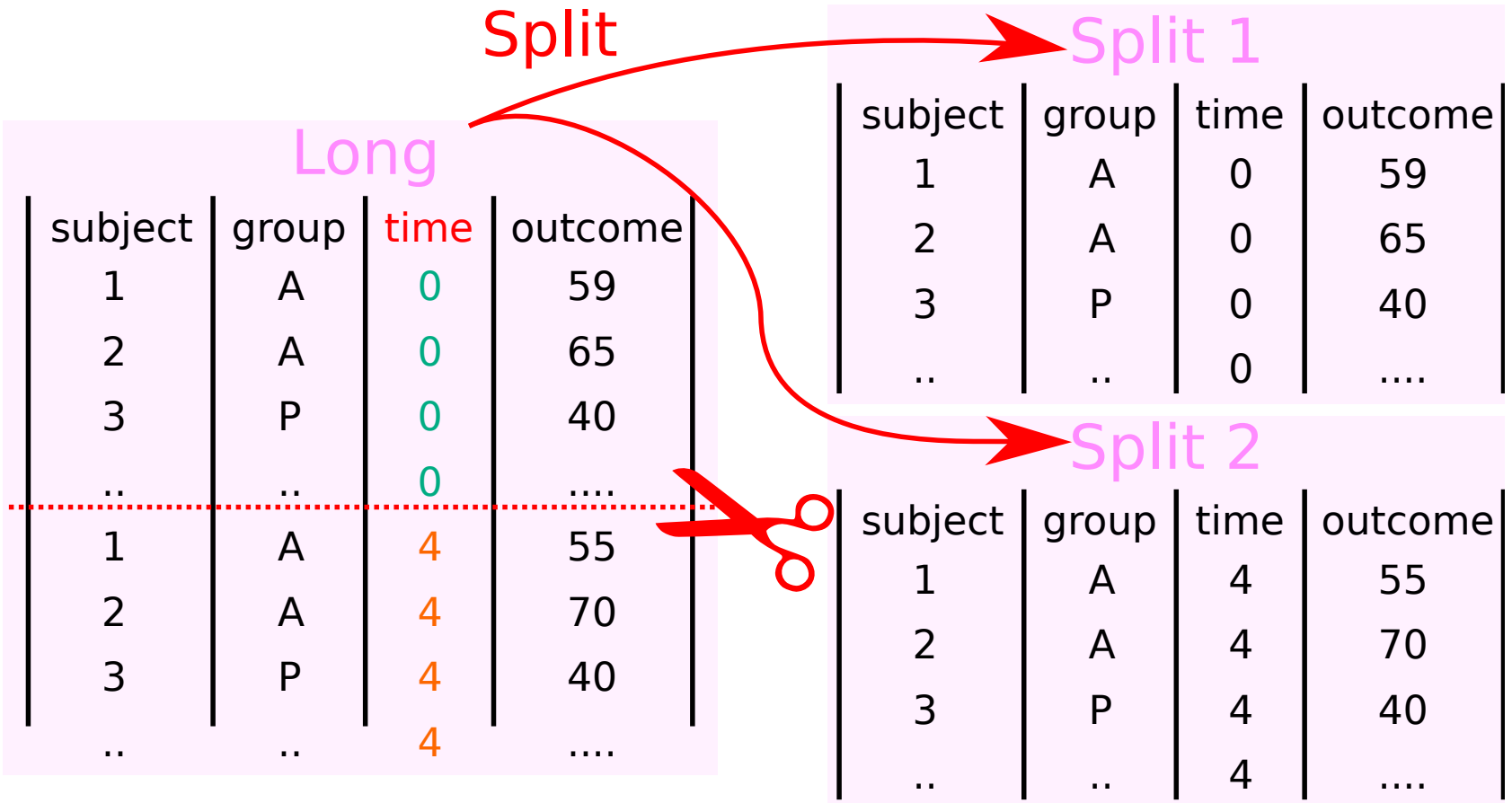


subject	group	time	outcome
1	A	0	59
1	A	4	55
2	A	0	65
2	A	4	70
3	P	0	40
3	P	4	40
..	..	..	....

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	....

Technical note: summary statistic in the long format

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



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	....

Split 1

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

Split 2

subject	group	time	outcome
1	A	4	55
2	A	4	70
3	P	4	40
..	..	4	....

Technical note: summary statistic in the long format

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

Split 1

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

Summarize

mean

59
65
40
....

Split 2

subject	group	time	outcome
1	A	4	55
2	A	4	70
3	P	4	40
..	..	4	.....

mean

55
70
40
....

Univariate approach

- adjustment resulting from working on change from baseline
- treatment effect assessment using a two sample t-test on the change
- pros and cons

General idea (1/2)

Define a (timepoint-specific) research question:

- e.g. treatment effect on the long-term change from baseline (here long-term means week 52)

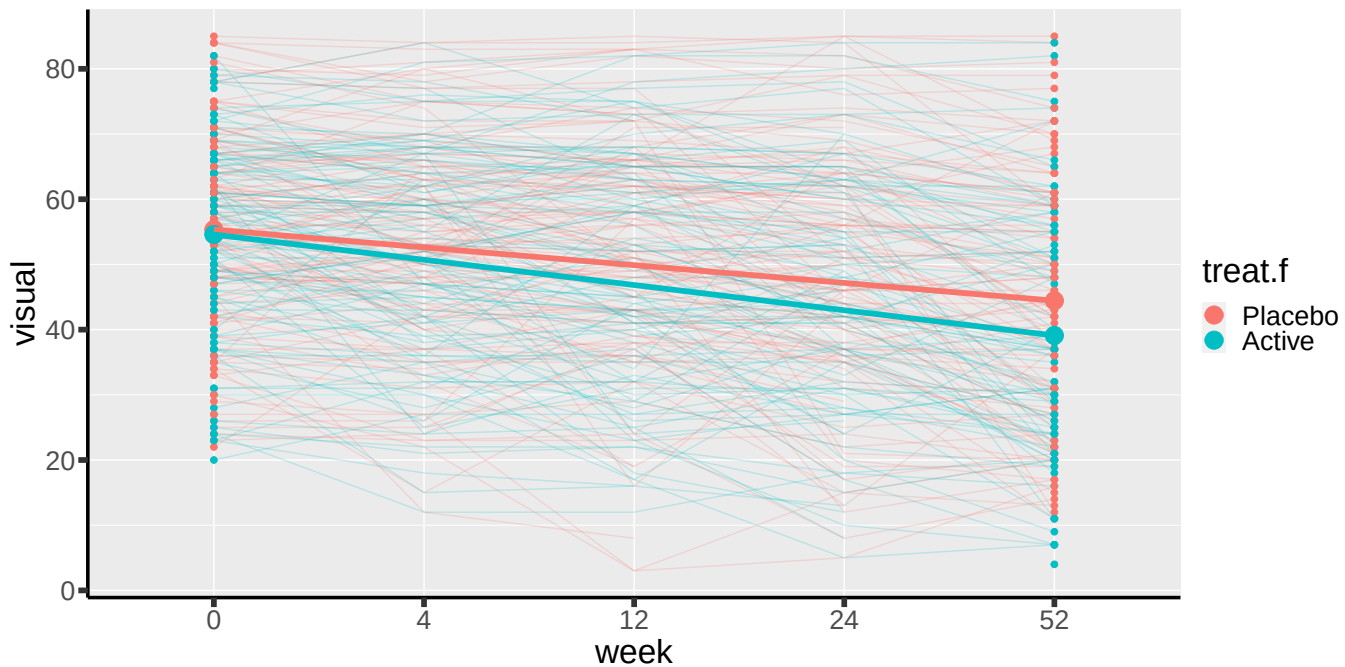
General idea (1/2)

Define a (timepoint-specific) research question:

- e.g. treatment effect on the long-term change from baseline (here long-term means week 52)

Restrict the dataset to the timepoint(s) of interest:

- e.g. week 0 and week 52



Why working on the change can be a good idea?

Describe the outcome value $Y_i(t)$ over of time t for individual i as:

$$Y_i(t) = \alpha(t) + \beta X_i + \varepsilon_i(t)$$

- $\alpha(t)$: time effect
- X_i: individual characteristics effects
  ⚠ assumed constant with constant effect
- unknown factors $\varepsilon_i(t)$ with variance σ^2
  (assumed constant over time to simplify)

The change in outcome between baseline and week 52 is:

Why working on the change can be a good idea?

Describe the outcome value $Y_i(t)$ over of time t for individual i as:

$$Y_i(t) = \alpha(t) + \beta X_i + \varepsilon_i(t)$$

- $\alpha(t)$: time effect
- X_i: individual characteristics effects
  ⚠ assumed constant with constant effect
- unknown factors $\varepsilon_i(t)$ with variance σ^2
  (assumed constant over time to simplify)

The change in outcome between baseline and week 52 is:

$$Y_i(52) - Y_i(0) = \alpha(52) - \alpha(0) + \varepsilon_i(52) - \varepsilon_i(0)$$

Why working on the change can be a good idea?

Describe the outcome value $Y_i(t)$ over of time t for individual i as:

$$Y_i(t) = \alpha(t) + \beta X_i + \varepsilon_i(t)$$

- $\alpha(t)$: time effect
- X_i: individual characteristics effects
  ⚠ assumed constant with constant effect
- unknown factors $\varepsilon_i(t)$ with variance σ^2
  (assumed constant over time to simplify)

The change in outcome between baseline and week 52 is:

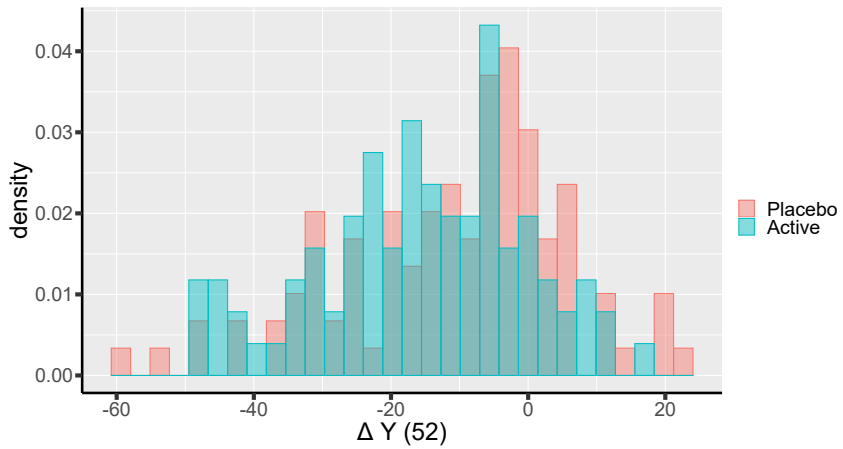
$$Y_i(52) - Y_i(0) = \alpha(52) - \alpha(0) + \varepsilon_i(52) - \varepsilon_i(0)$$

When $\rho > 1/2$,
$\mathbb{V}ar\left[Y_i(52) - Y_i(0)\right] = 2\sigma^2(1-\rho)$ is smaller than $\mathbb{V}ar\left[Y_i(t)\right]$.

General idea (2/2)

Restrict the dataset to the timepoint(s) of interest

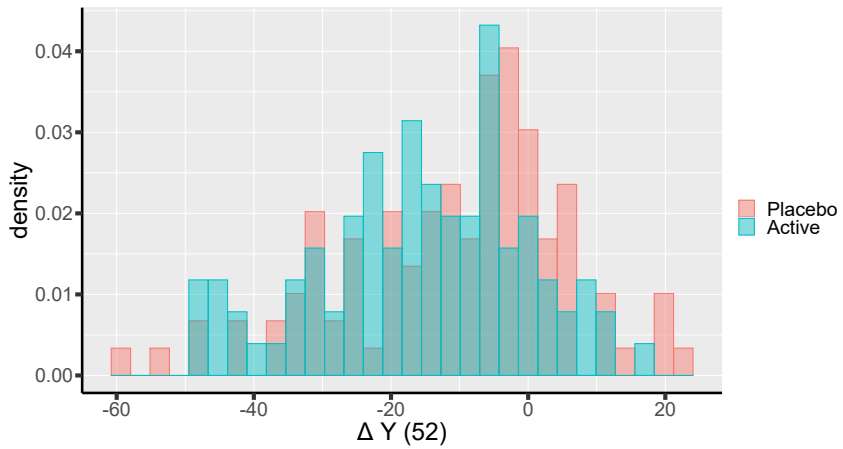
$\rightarrow$ compute the change in outcome: $\Delta Y(52) = Y(52) - Y(0)$



General idea (2/2)

Restrict the dataset to the timepoint(s) of interest

$\rightarrow$ compute the change in outcome: $\Delta Y(52) = Y(52) - Y(0)$



Use an appropriate statistical test to assess the treatment effect:

- e.g. a t-test is optimal when, within groups, the outcome is normally distributed:

$$\Delta Y(52)|G = 0 \sim \mathcal{N}(\mu_0, \tau_0)$$
$$\Delta Y(52)|G = 1 \sim \mathcal{N}(\mu_1, \tau_1)$$

Is it a good idea? (1/2)

✔ makes no assumption about the treatment effect over time

✔ simple to carry out

✔ naturally accounts for some covariates, even when unobserved.

Is it a good idea? (1/2)

✔ makes no assumption about the treatment effect over time

✔ simple to carry out

⚠ . . . except in presence of missing values!
in the previous slides, complete case analysis was performed which is biased in presence of informative censoring.

✔ naturally accounts for some covariates, even when unobserved.

Is it a good idea? (1/2)

✔ makes no assumption about the treatment effect over time

✔ simple to carry out

⚠ . . . except in presence of missing values!
in the previous slides, complete case analysis was performed which is biased in presence of informative censoring.

✔ naturally accounts for some covariates, even when unobserved.

⚠ does not account for unbalanced in baseline score
which can lead to bias if baseline score is correlated to change (Vickers and Altman, 2001).

→ use a linear model instead $Y_i(52) = \alpha + \beta X_i + \gamma Y_i(0) + \varepsilon_i(52)$

Is it a good idea? (2/2)

When looking at several timepoints:

	dmean in Placebo	dmean in Active	difference	p.value
week 4	-1.30	-3.51	-2.21	0.04
week 12	-2.27	-5.88	-3.61	0.02
week 24	-5.71	-9.07	-3.36	0.08
week 52	-11.18	-15.48	-4.30	0.06

Is it a good idea? (2/2)

When looking at several timepoints:

	dmean in Placebo	dmean in Active	difference	p.value
week 4	-1.30	-3.51	-2.21	0.04
week 12	-2.27	-5.88	-3.61	0.02
week 24	-5.71	-9.07	-3.36	0.08
week 52	-11.18	-15.48	-4.30	0.06


multiple testing issue

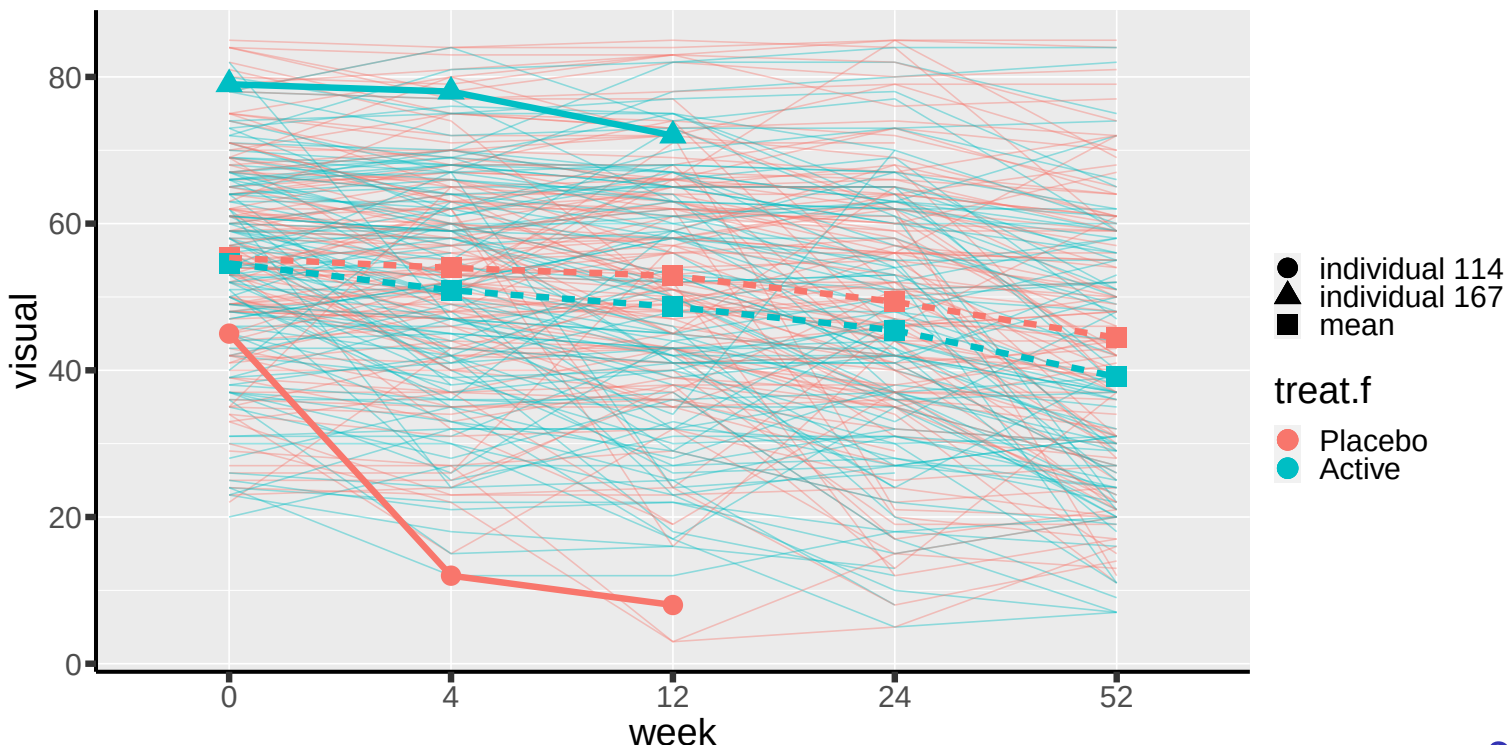

estimates are timepoint-specific: what about week 30?

Multivariate approach

- intuition behind handling missing values using a multivariate model
- parametrization of a linear mixed model
- pros and cons

Handling missing value: full information

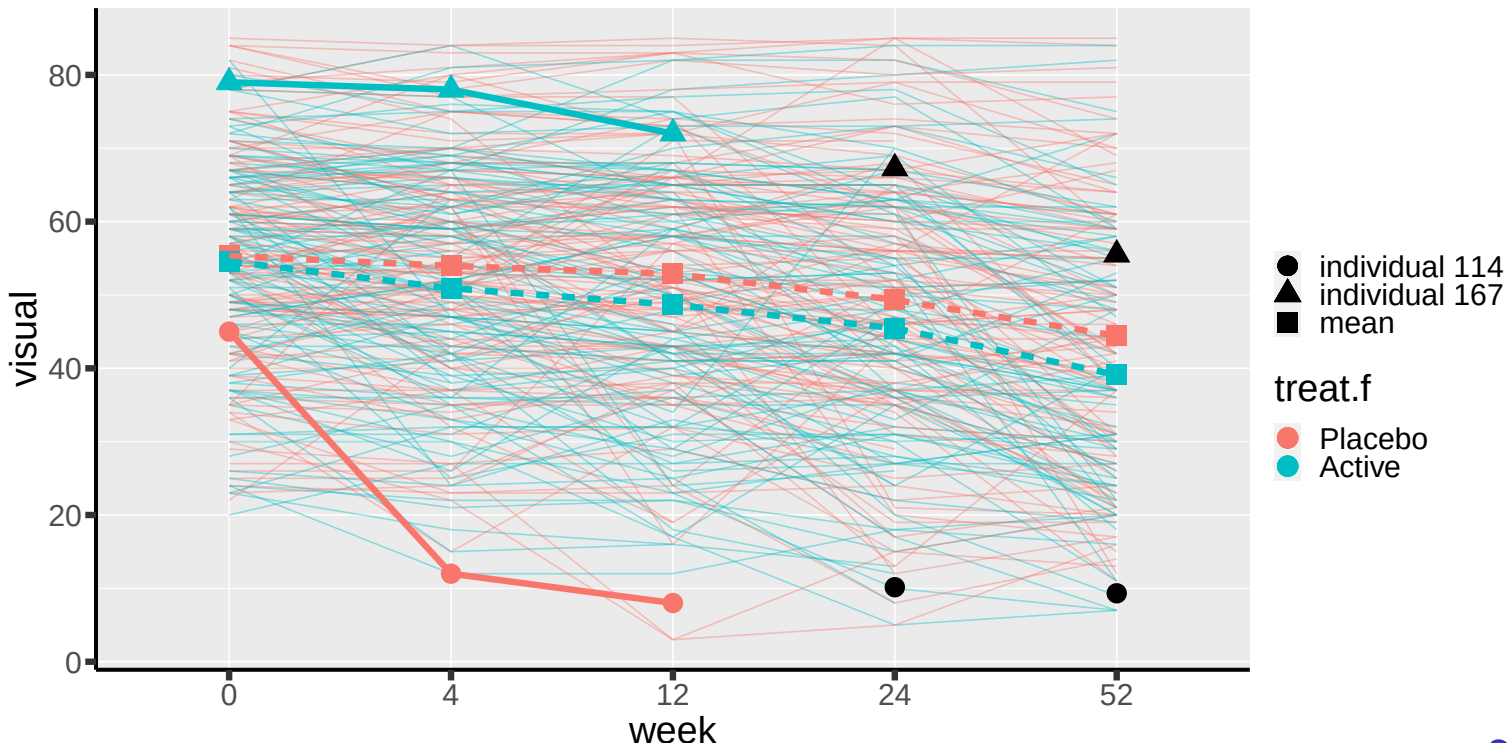
Previously, individuals 114 and 167 were removed from the analysis because of missing outcome at week 52. Can we do better?



Handling missing value: full information

Previously, individuals 114 and 167 were removed from the analysis because of missing outcome at week 52. Can we do better?

Yes! Using (i) their observed timepoints
(ii) the modeled mean/covariance
to "guess" their outcome at week 52



Is it a good idea?

✔ the mean will be robust to drop-out depending on past outcome values.

(not the case when using complete case analysis)

✔ the estimation of the mean will be more precise.

However it requires a more complex model:

- more difficult analysis to carry-out

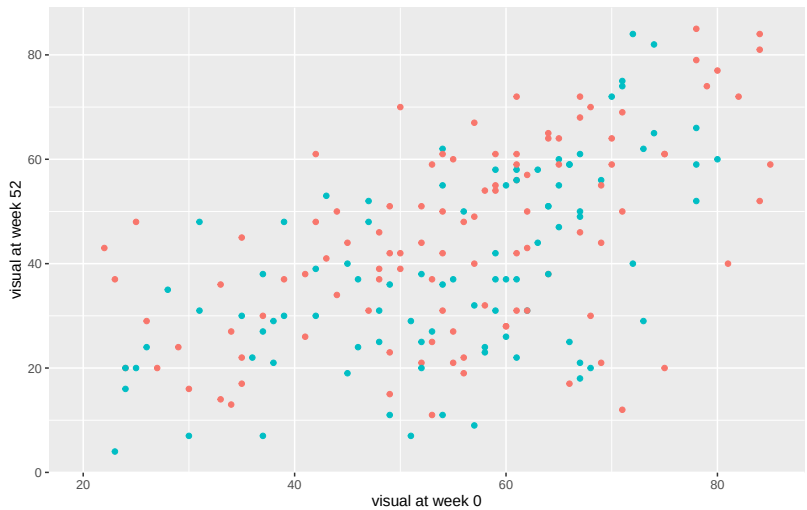
⚠ extra-hypotheses,
e.g. linear association between week 0 and 52 values

How can we do that?

Using a linear model relating :

- outcomes at timepoints where the subject has data
- at timepoints where the subject has no data

The relationship is estimated using data from the other subjects.

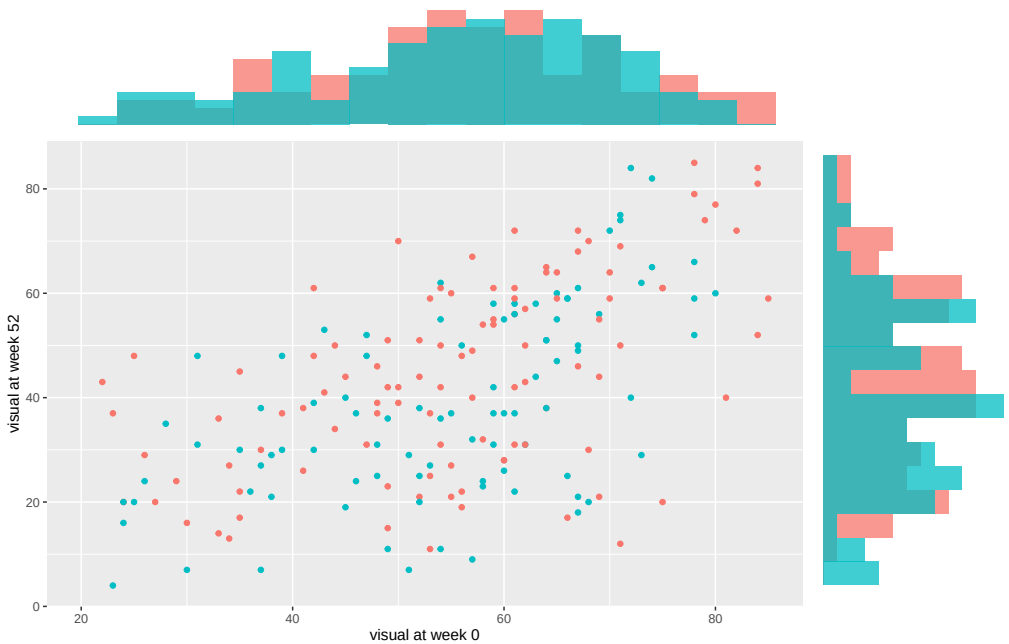


How can we do that?

Using a linear model relating :

- outcomes at timepoints where the subject has data
- at timepoints where the subject has no data

The relationship is estimated using data from the other subjects.

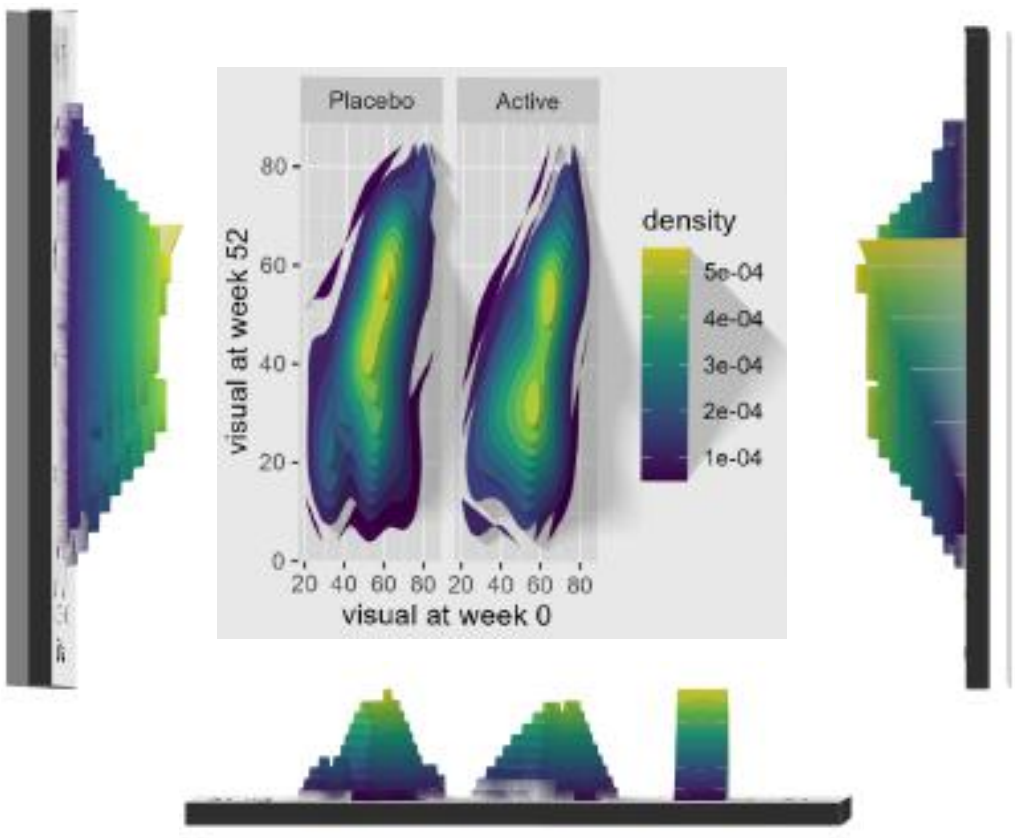


How can we do that?

Using a linear model relating :

- outcomes at timepoints where the subject has data
- at timepoints where the subject has no data

The relationship is estimated using data from the other subjects.

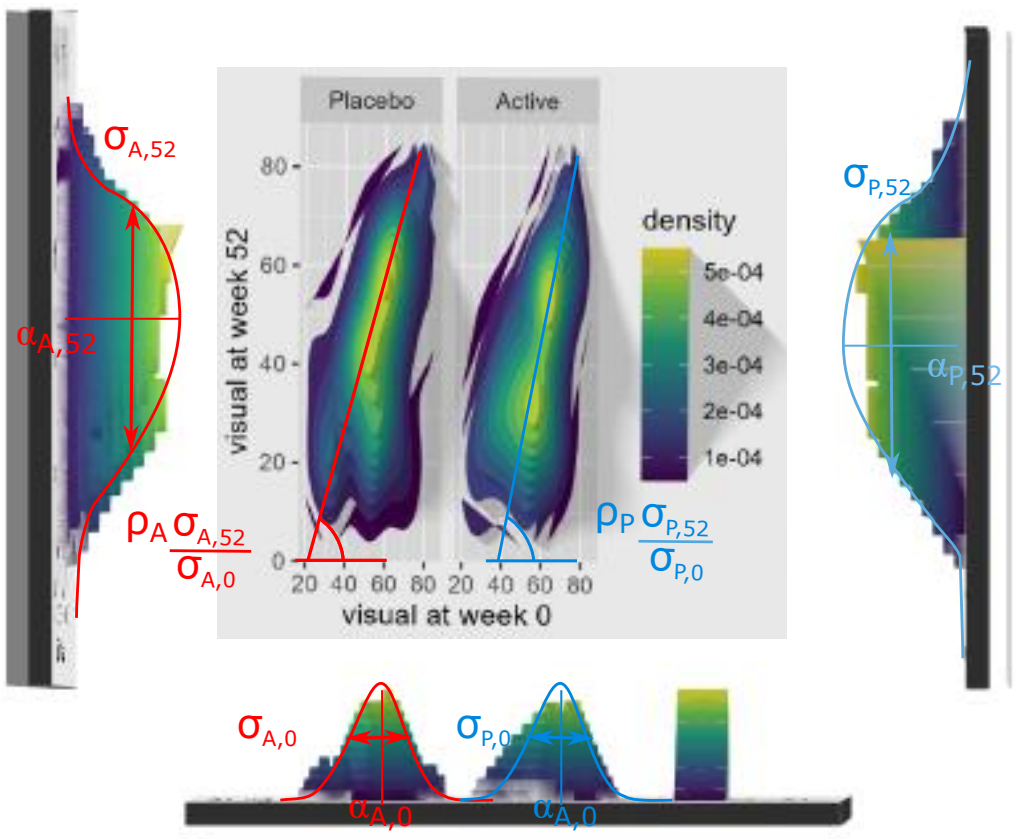


How can we do that?

Using a linear model relating :

- outcomes at timepoints where the subject has data
- at timepoints where the subject has no data

The relationship is estimated using data from the other subjects.



How can we do that?

Using a linear model relating :

- outcomes at timepoints where the subject has data
- at timepoints where the subject has no data

The relationship is estimated using data from the other subjects.

Formally, the expected value at one timepoint given the observed value at another is:

(assuming a joint normal distribution over time)

$$\widehat{Y}_i(52) = \alpha(52) + \rho(0,52)\frac{\sigma(52)}{\sigma(0)}\left(Y_i(0) - \alpha(0)\right)$$

$\rightarrow$ we need not only to model the **mean** but also the **variance** and **correlation** over time!

Implementation

In practice we will use what is called a mixed **mixed model**:

- generalization of the univariate linear model (`lm` in **R**)
- need more inputs: variance and correlation structure
- 😕 format of these "new" inputs is software dependent

There are several **R** package implementing mixed models:

- nlme and lme4: recommended ones
- LMMstar: narrower scope but should be more user-friendly

Example in R with 2 timepoints

```
library(LMMstar)

e.lmm <- lmm(visual ~ treat.f*week, ## mean structure
      repetition = ~ week | subject,
      structure = "UN", ## variance/correlation structure
      data = armd.long[armd.long$week %in% c("0","52"),])

model.tables(e.lmm)
```

	estimate	se	df	lower	upper	p.value
(Intercept)	55.336	1.37	238	52.64	58.029	0.00e+00
treat.fActive	-0.758	1.93	238	-4.55	3.035	6.94e-01
week52	-11.095	1.55	196	-14.15	-8.038	1.61e-11
treat.fActive:week52	-4.383	2.27	198	-8.87	0.103	5.54e-02

Example in R with 2 timepoints

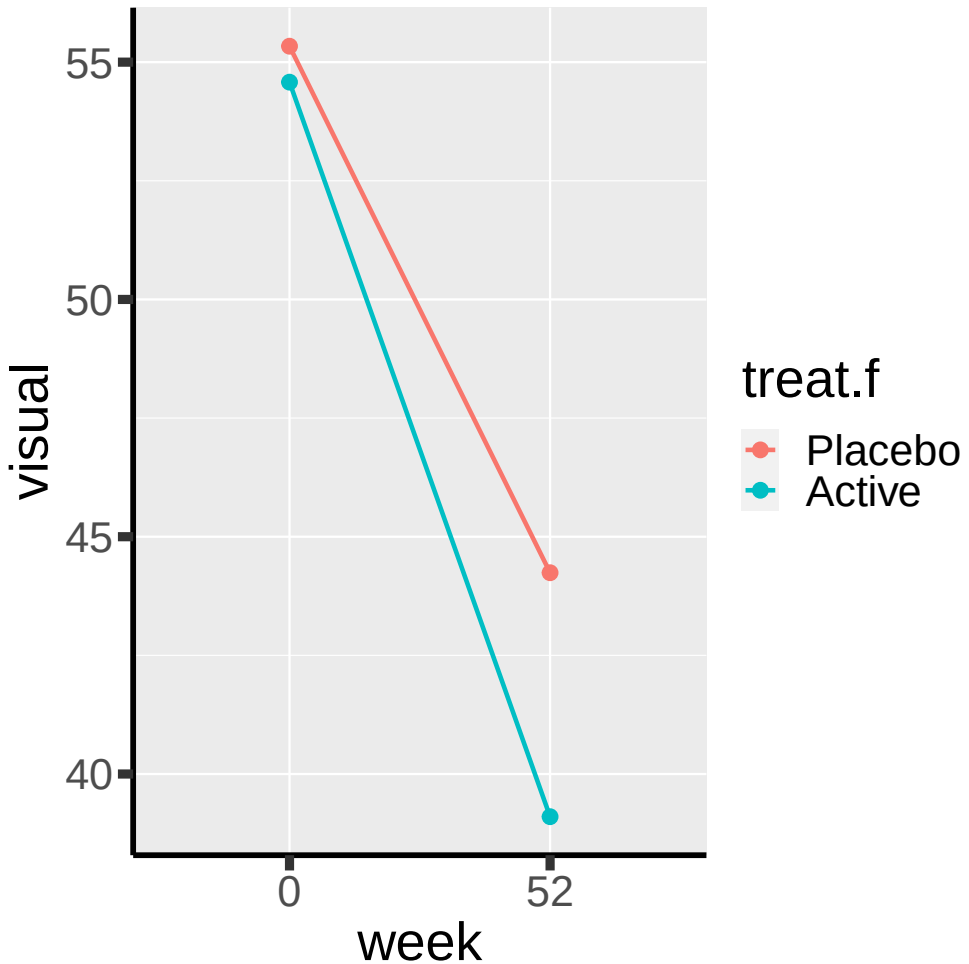
```
library(LMMstar)

e.lmm <- lmm(visual ~ treat.f*week, ## mean structure
      repetition = ~ week | subject,
      structure = "UN", ## variance/correlation structure
      data = armd.long[armd.long$week %in% c("0","52"),])

model.tables(e.lmm)
```

	estimate	se	df	lower	upper	p.value
(Intercept)	55.336	1.37	238	52.64	58.029	0.00e+00
treat.fActive	-0.758	1.93	238	-4.55	3.035	6.94e-01
week52	-11.095	1.55	196	-14.15	-8.038	1.61e-11
treat.fActive:week52	-4.383	2.27	198	-8.87	0.103	5.54e-02

What are those coefficients?

Let's look at the fitted values (1/2)

```
plot(e.lmm, type = "fit", ci = FALSE)
```

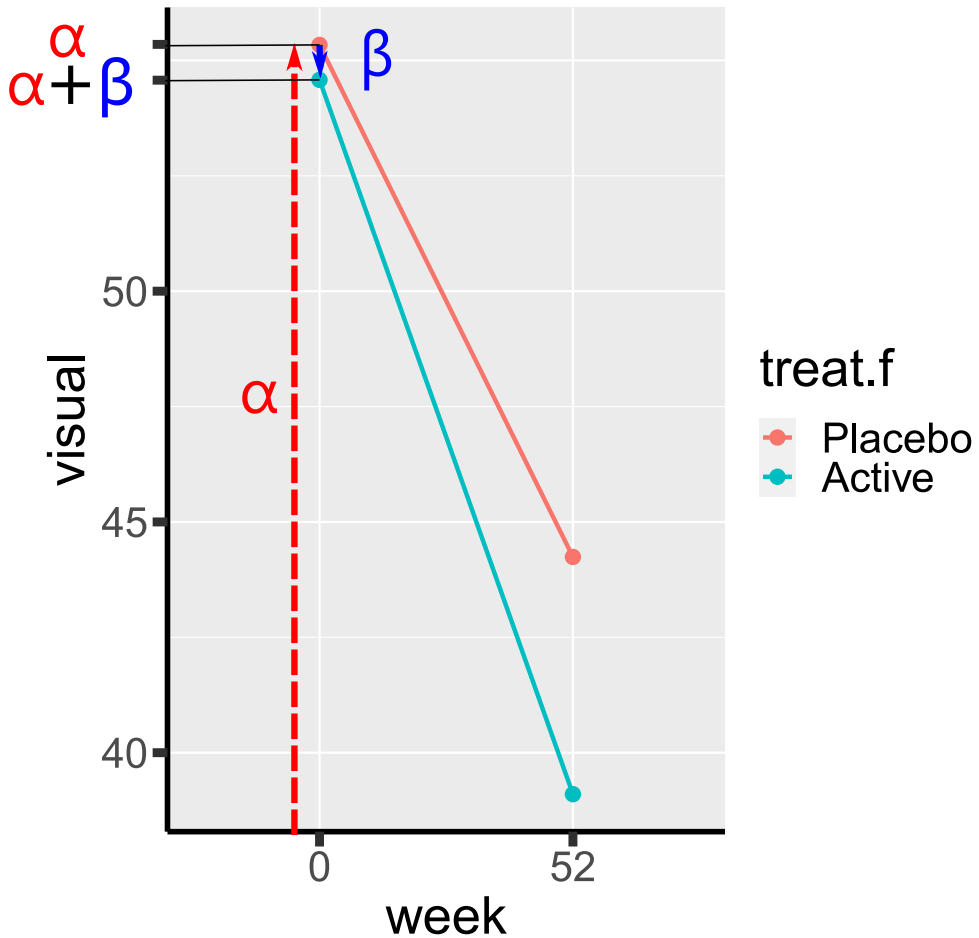



```
coef(e.lmm)
```

```
(Intercept)
55.3361345
treat.fActive
-0.7576221
week52
-11.0948777
treat.fActive:week52
-4.3831236
```

Let's look at the fitted values (1/2)

```
plot(e.lmm, type = "fit", ci = FALSE)
```

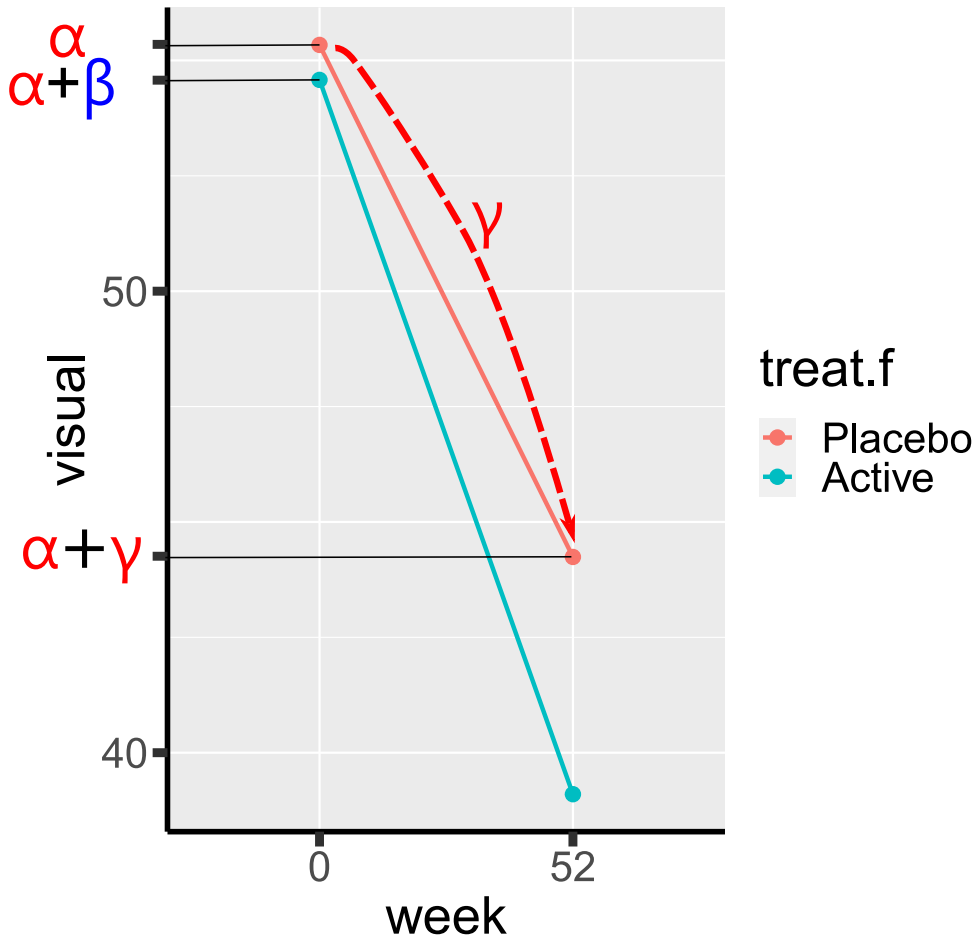



```
coef(e.lmm)
```

```
(Intercept)
55.3361345
treat.fActive
-0.7576221
week52
-11.0948777
treat.fActive:week52
-4.3831236
```

Let's look at the fitted values (1/2)

```
plot(e.lmm, type = "fit", ci = FALSE)
```

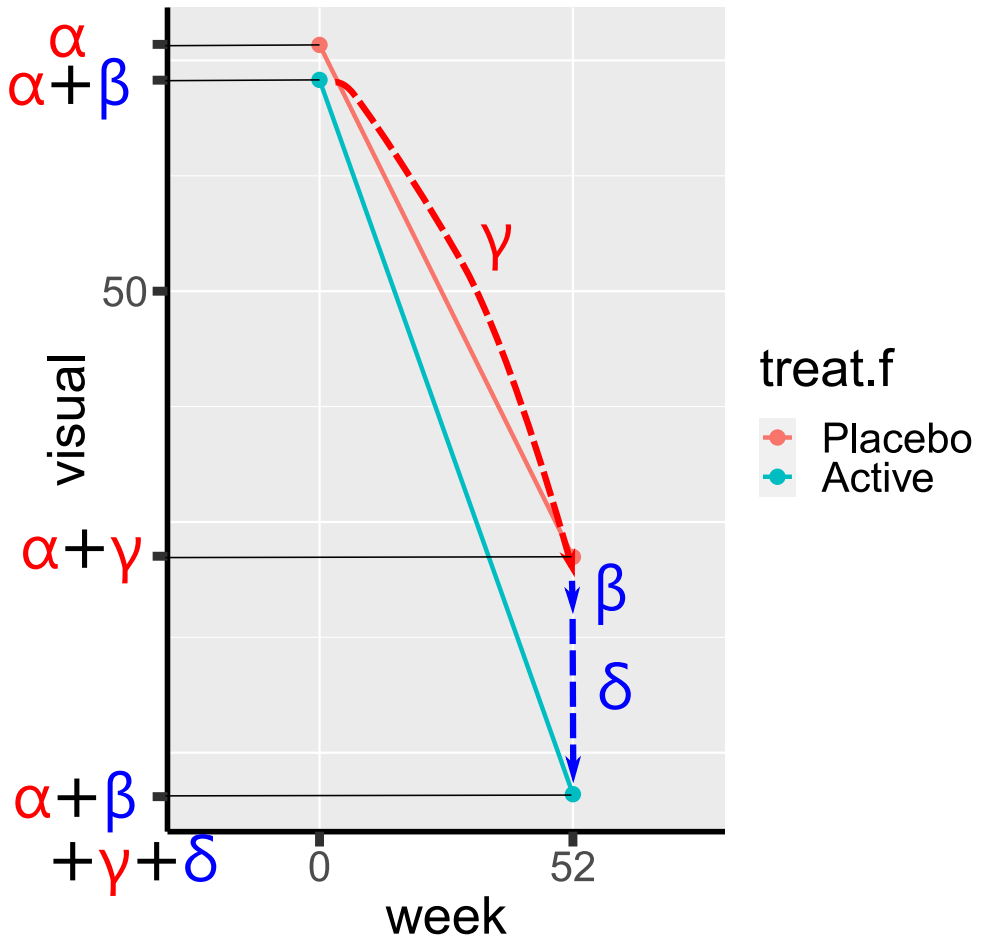



```
coef(e.lmm)
```

```
(Intercept)
55.3361345
treat.fActive
-0.7576221
week52
-11.0948777
treat.fActive:week52
-4.3831236
```

Let's look at the fitted values (1/2)

```
plot(e.lmm, type = "fit", ci = FALSE)
```

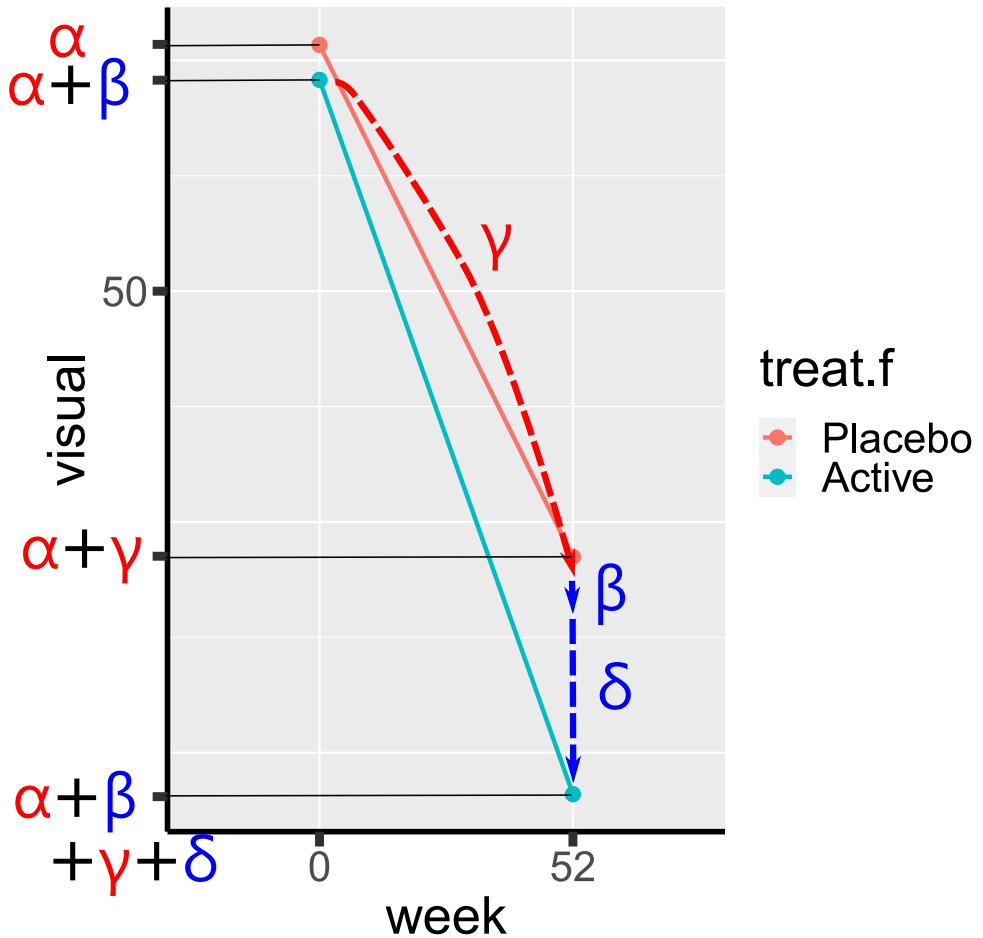



```
coef(e.lmm)
```

```
(Intercept)
55.3361345
treat.fActive
-0.7576221
week52
-11.0948777
treat.fActive:week52
-4.3831236
```

Let's look at the fitted values (2/2)

```
plot(e.lmm, type = "fit", ci = FALSE)
```




```
dummy.coef(e.lmm)
     [,1:3]
```

	treat.f	week	estimate
1	Placebo	0	55.33613
2	Active	0	54.57851
3	Placebo	52	44.24126
4	Active	52	39.10051

Parametrization 2 timepoints (CS)

Compound symmetry or random intercept structure:

$$\begin{matrix}\text{placebo} \\ \text{group}\end{matrix} \begin{bmatrix} Y_0 \\ Y_{52} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \alpha \\ \alpha+\gamma \end{bmatrix}, \sigma_0^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

$$\begin{matrix}\text{active} \\ \text{group}\end{matrix} \begin{bmatrix} Y_0 \\ Y_{52} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \alpha+\beta \\ \alpha+\beta+\gamma+\delta \end{bmatrix}, \sigma_0^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

Parametrization 2 timepoints (CS)

Compound symmetry or random intercept structure:

$$\begin{matrix}\text{placebo}\\ \text{group}\end{matrix}\begin{bmatrix}Y_0\\ Y_{52}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\alpha\\ \alpha+\gamma\end{bmatrix}, \sigma_0^2\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$$

$$\begin{matrix}\text{active}\\ \text{group}\end{matrix}\begin{bmatrix}Y_0\\ Y_{52}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\alpha+\beta\\ \alpha+\beta+\gamma+\delta\end{bmatrix}, \sigma_0^2\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$$

⚠ we assume no treatment **and** no time effect on the variance/correlation

Parametrization 2 timepoints (UN)

Unstructured variance/correlation:

$$\begin{matrix}\text{placebo}\\ \text{group}\end{matrix}\begin{bmatrix}Y_0\\ Y_{52}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\alpha\\ \alpha+\gamma\end{bmatrix}, \sigma_0^2\begin{bmatrix}1 & \rho k_{52}\\ \rho k_{52} & k_{52}^2\end{bmatrix}\right)$$

$$\begin{matrix}\text{active}\\ \text{group}\end{matrix}\begin{bmatrix}Y_0\\ Y_{52}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\alpha+\beta\\ \alpha+\beta+\gamma+\delta\end{bmatrix}, \sigma_0^2\begin{bmatrix}1 & \rho k_{52}\\ \rho k_{52} & k_{52}^2\end{bmatrix}\right)$$

```
coef(e.lmm, effects = c("variance","correlation"))
```

```
     sigma        k.52  rho(0,52)
14.9115119  1.2397277  0.5612167
```

Parametrization 2 timepoints (UN)

Unstructured variance/correlation:

$$\begin{matrix}\text{placebo}\\ \text{group}\end{matrix}\begin{bmatrix}Y_0\\ Y_{52}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\alpha\\ \alpha+\gamma\end{bmatrix}, \sigma_0^2\begin{bmatrix}1 & \rho k_{52}\\ \rho k_{52} & k_{52}^2\end{bmatrix}\right)$$

$$\begin{matrix}\text{active}\\ \text{group}\end{matrix}\begin{bmatrix}Y_0\\ Y_{52}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\alpha+\beta\\ \alpha+\beta+\gamma+\delta\end{bmatrix}, \sigma_0^2\begin{bmatrix}1 & \rho k_{52}\\ \rho k_{52} & k_{52}^2\end{bmatrix}\right)$$

⚠ we assume no treatment effect on the variance/correlation

```
coef(e.lmm, effects = c("variance","correlation"))
```

```
     sigma        k.52  rho(0,52)
14.9115119  1.2397277  0.5612167
```

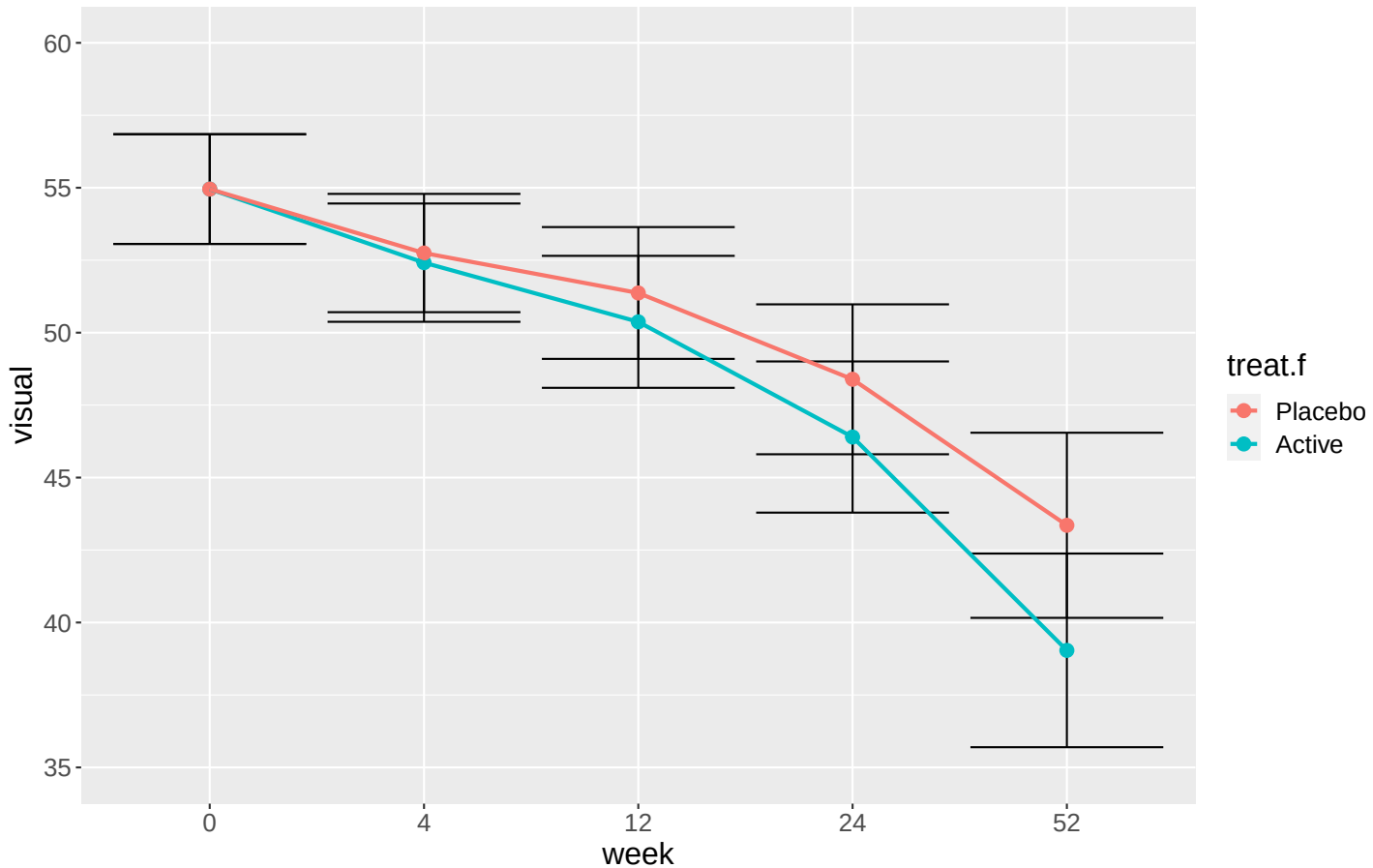
Treatment effect proportional to duration

```
## week     : categorical variable ("week0", "week4", ...)
## week.num: numeric variable (0, 4, ...)
eLin.lmm <- lmm(visual ~ 0 + week + week.num:treat.f,
  repetition = ~ week | subject,
  structure = "UN",
  data = armd.long)

model.tables(eLin.lmm)
```

	estimate	se	df	lower	upper	p.value
week0	54.954	0.9608	239	53.061	56.84694	0.0000
week4	52.748	1.0359	240	50.707	54.78821	0.0000
week12	51.369	1.1544	257	49.096	53.64265	0.0000
week24	48.391	1.3141	281	45.804	50.97755	0.0000
week52	43.353	1.6209	232	40.160	46.54704	0.0000
week.num:treat.fActive	-0.083	0.0409	187	-0.164	-0.00231	0.0439

Visualisation

```
autoplot(eLin.lmm)
```



Warp-up

Why using mixed models?

Generalize t-test on the change:

- equivalent with 2 endpoints and no missing data

Better handling of missing values:

- full information instead of complete case analysis

✔ no need to model the cause of censoring

⚠ require valid model for the mean/covariance structure

Can ease interpretability:

- imposing constant or linear treatment effect over time

When not to use mixed models?

No missing data and only two timepoints

- a univariate analysis on the change from baseline is often enough

Very small sample size:

- model parameters can be difficult to estimate
- possible influation of type 1 error with the "usual" tests

In presence of competing risks (e.g. death)

- mixed model are not a "magic" solution for missing values ...

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

What we have seen today

✔ Introduction to repeated measurements

- definition and examples of study design
- challenges for the statistical analysis

✔ Example of longitudinal study

- descriptive statistics and plots for repeated measurements
- concerns due to the presence of missing values
- what is a long and wide format

✔ Univariate approach

- adjustment resulting from working on change from baseline
- treatment effect assessment using a two sample t-test on the change
- pros and cons

✔ Multivariate approach

- intuition behind handling missing values using a multivariate model
- parametrization of a linear mixed model
- pros and cons

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

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.

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.

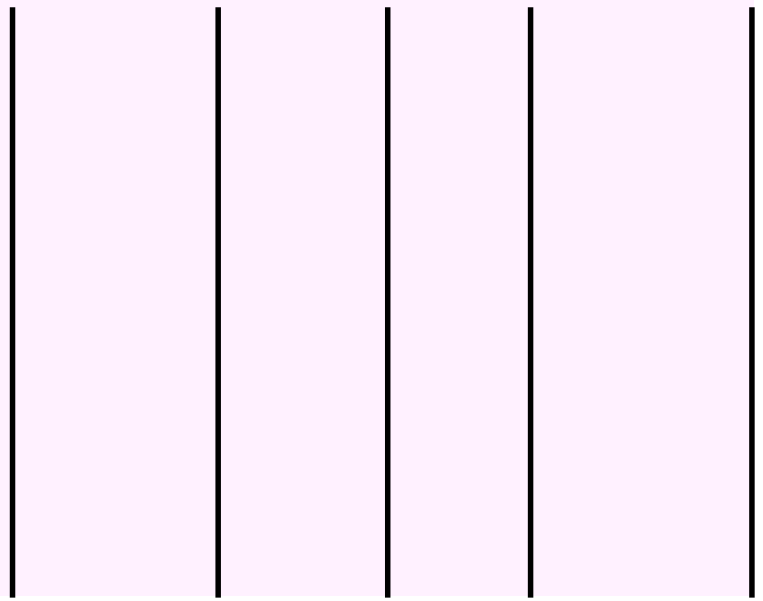
Vickers, A. J. and Altman, D. G. (2001). Analysing controlled trials with baseline and follow up measurements. *Bmj*, 323(7321):1123–1124.

Wide to long format

Wide

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

Long



Wide to long format

Wide

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

Long

subject	group	time	outcome
		0	
		4	
		0	
		4	
		0	
		4	
		..	

Wide to long format

Wide

subject	group	outcome (time = 0)	outcome (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	
		..	

Equivalence t-test and mixed model (1/3)

t-test (complete case week 0 and 52):

```
armd.wideCC <- na.omit(armd.wide[,c("subject","treat.f","
    visual0","visual52")])
t.test(visual52-visual0 ~ treat.f,
       data = armd.wideCC)
```

```
        Welch Two Sample t-test

data:  visual52 - visual0 by treat.f
t = 1.8842, df = 191.47, p-value = 0.06106
alternative hypothesis: true difference in means between group P
95 percent confidence interval:
 -0.2013017  8.7949525
sample estimates:
mean in group Placebo  mean in group Active
            -11.18095              -15.47778
```

Equivalence t-test and mixed model (2/3)

Mixed model on the change (complete case week 0 and 52):

```
armd.wideCC$change <- armd.wideCC$visual52-armd.wideCC$
    visual0
e2CC.lmm <- lmm(change ~ treat.f,
  repetition = ~ treat.f | subject, structure = "UN",
  data = armd.wideCC)

model.tables(e2CC.lmm)
```

```
                estimate   se  df  lower  upper  p.value
(Intercept)        -11.2 1.60 104 -14.36 -8.002 2.94e-10
treat.fActive       -4.3 2.28 192  -8.79  0.201 6.11e-02
```

Equivalence t-test and mixed model (3/3)

Mixed model on time-specific data (complete case week 0 and 52):

```
test.CC <- armd.long$week %in% c("0","52") & armd.long$
    subject %in% subjCC
e2CC.lmm <- lmm(visual ~ week*treat.f,
  repetition = ~ week | subject, structure = "UN",
  data = armd.long[testCC,])

model.tables(e2CC.lmm)["week52:treat.fActive",,drop=FALSE]
c("Placebo" = as.double(coef(e2CC.lmm)["week52"]),
  "Active" = sum(coef(e2CC.lmm)[c("week52","week52:treat.
    fActive")]))
```

```
                      estimate   se  df lower upper p.value
week52:treat.fActive     -4.3 2.29 193 -8.82 0.224  0.0624
  Placebo    Active
-11.18095 -15.47778
```