Statistical analysis of repeated measurements
Lecture 3: random effects models

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

[1] Neurobiology Research Unit, University Hospital of Copenhagen, Rigshospitalet.

[2] Section of Biostatistics, Department of Public Health, University of Copenhagen.

29-11-2022

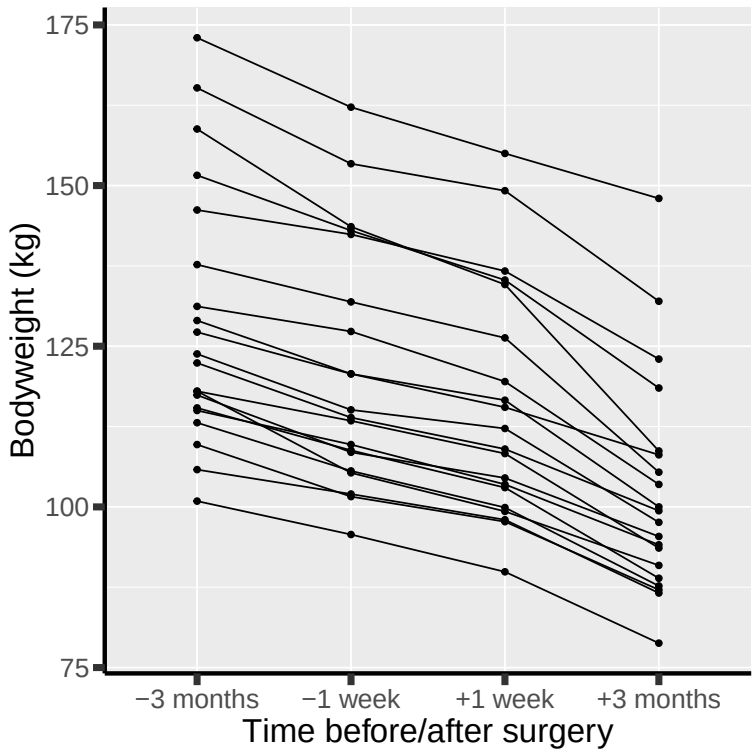
The course so far

Lecture 1 & 2: one arm and two arms follow-up studies

- describe and compare the change over time using LMM
- confirmatory clinical study

LMM: linear mixed model

- linear model for the mean
  categorical vs. continuous time
  baseline adjustment
- model for the variance
  compound symmetry,
  unstructured,
  stratified unstructured
- can handle missing data



Mixed model as a Gaussian model

Under normality assumption, the data $\mathbf{Y} = (Y_1, \ldots, Y_t)$ is fully characterized by:

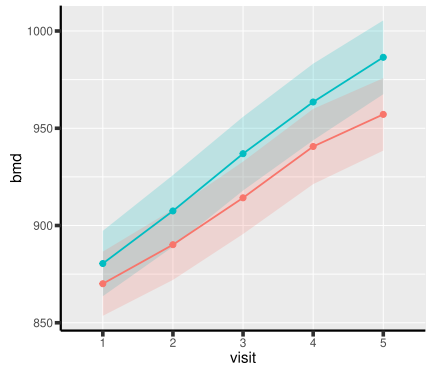
- a mean vector $\mu = (\mu_1, \ldots, \mu_t)$
- a covariance matrix Ω

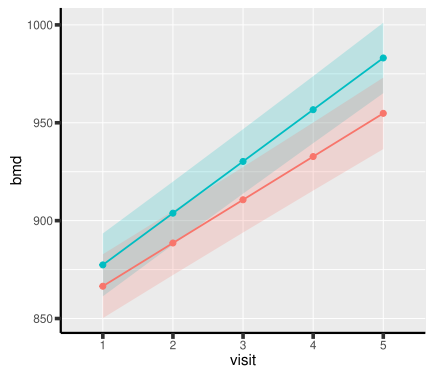
(i.e. standard deviations $\boldsymbol{\sigma}$ and correlations $\boldsymbol{\rho}$)

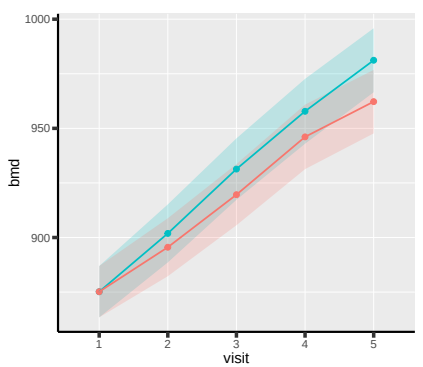
$$\mathbf{Y} \sim \mathcal{N}(\mu, \Omega)$$

The mixed model is a way to **parametrize** a multivariate normal distribution and **estimate** these parameters.

Examples of mean structures



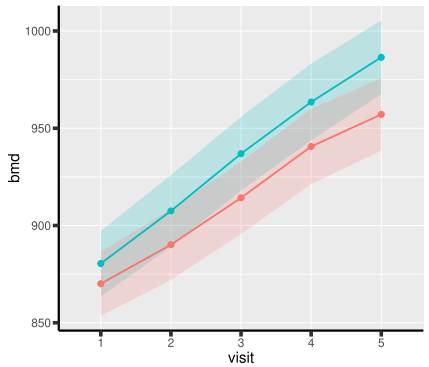


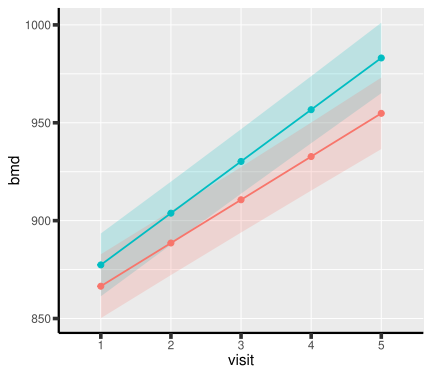


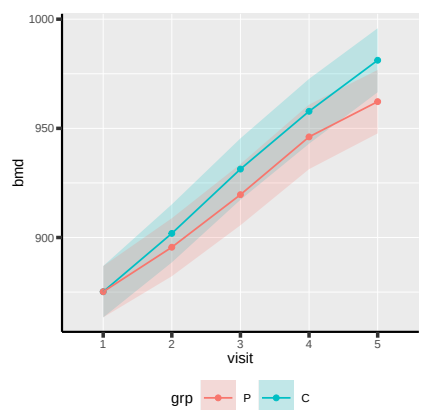
Examples of mean structures

Time as categorical variable

time	Placebo (E=0)	Calcium (E=1)
baseline	α	$\alpha + \gamma$
6 months	$\alpha + \beta_1$	$\alpha + \beta_1 + \gamma + \delta_1$
12 months	$\alpha + \beta_2$	$\alpha + \beta_2 + \gamma + \delta_2$
18 months	$\alpha + \beta_3$	$\alpha + \beta_3 + \gamma + \delta_3$
24 months	$\alpha + \beta_4$	$\alpha + \beta_4 + \gamma + \delta_4$



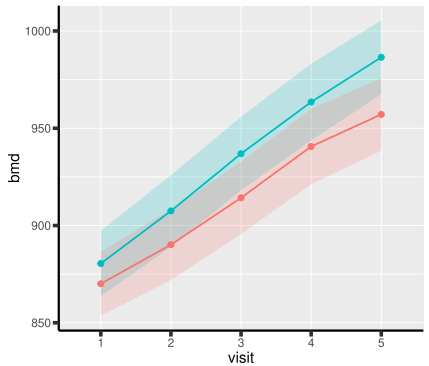




Examples of mean structures

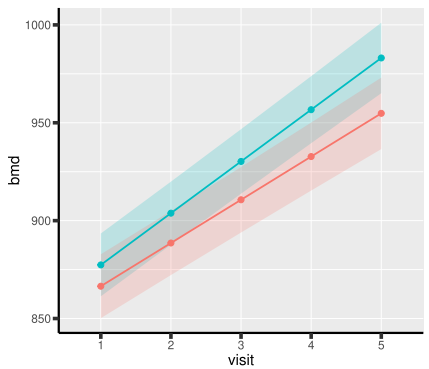
Time as categorical variable

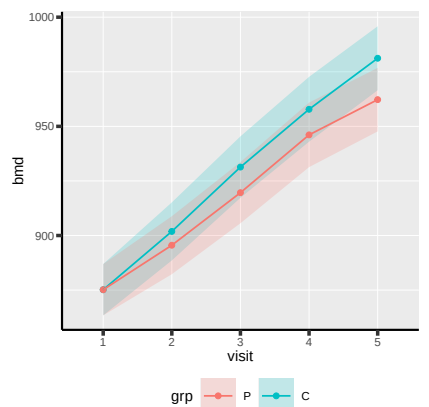
time	Placebo (E=0)	Calcium (E=1)
baseline	α	$\alpha + \gamma$
6 months	$\alpha + \beta_1$	$\alpha + \beta_1 + \gamma + \delta_1$
12 months	$\alpha + \beta_2$	$\alpha + \beta_2 + \gamma + \delta_2$
18 months	$\alpha + \beta_3$	$\alpha + \beta_3 + \gamma + \delta_3$
24 months	$\alpha + \beta_4$	$\alpha + \beta_4 + \gamma + \delta_4$



Time as numeric variable

time	Placebo (E=0)	Calcium (E=1)
baseline	α	$\alpha + \gamma$
6 months	$\alpha + 0.5\beta$	$\alpha + 0.5\beta + \gamma + 0.5\delta$
12 months	$\alpha + \beta$	$\alpha + \beta + \gamma + \delta$
18 months	$\alpha + 1.5\beta$	$\alpha + 1.5\beta + \gamma + 1.5\delta$
24 months	$\alpha + 2\beta$	$\alpha + 2\beta + \gamma + 2\delta$

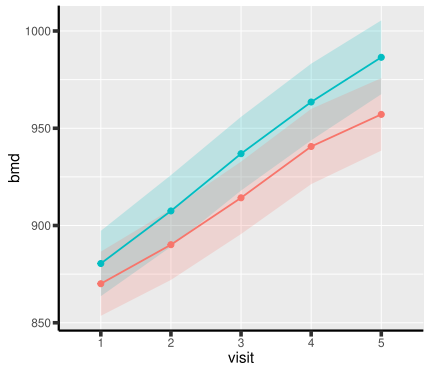




Examples of mean structures

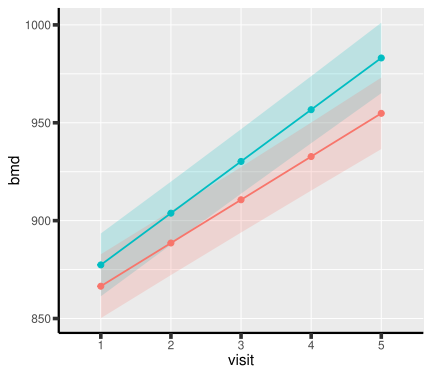
Time as categorical variable

time	Placebo (E=0)	Calcium (E=1)
baseline	α	$\alpha + \gamma$
6 months	$\alpha + \beta_1$	$\alpha + \beta_1 + \gamma + \delta_1$
12 months	$\alpha + \beta_2$	$\alpha + \beta_2 + \gamma + \delta_2$
18 months	$\alpha + \beta_3$	$\alpha + \beta_3 + \gamma + \delta_3$
24 months	$\alpha + \beta_4$	$\alpha + \beta_4 + \gamma + \delta_4$



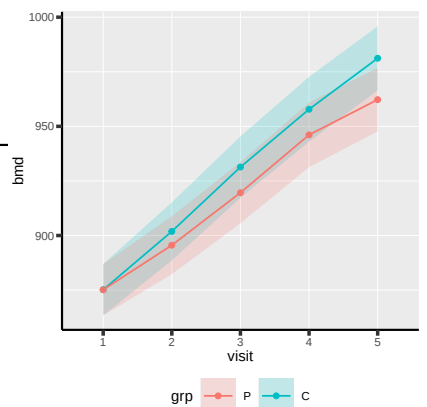
Time as numeric variable

time	Placebo (E=0)	Calcium (E=1)
baseline	α	$\alpha + \gamma$
6 months	$\alpha + 0.5\beta$	$\alpha + 0.5\beta + \gamma + 0.5\delta$
12 months	$\alpha + \beta$	$\alpha + \beta + \gamma + \delta$
18 months	$\alpha + 1.5\beta$	$\alpha + 1.5\beta + \gamma + 1.5\delta$
24 months	$\alpha + 2\beta$	$\alpha + 2\beta + \gamma + 2\delta$



Time as categorical variable + baseline adjustment

time	Placebo (E=0)	Calcium (E=1)
baseline	$\mu(0,0) = \alpha$	$\mu(0,1) = \alpha$
6 months	$\mu(0.5,0) = \alpha + \beta_1$	$\mu(0.5,1) = \alpha + \beta_1 + \delta_1$
12 months	$\mu(1,0) = \alpha + \beta_2$	$\mu(1,1) = \alpha + \beta_2 + \delta_2$
18 months	$\mu(1.5,0) = \alpha + \beta_3$	$\mu(1.5,1) = \alpha + \beta_3 + \delta_3$
24 months	$\mu(2,0) = \alpha + \beta_4$	$\mu(2,1) = \alpha + \beta_4 + \delta_4$



Examples of correlation structures

$$\begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \rho_{1,2} & \rho_{1,3} & \rho_{1,4} \\ \rho_{1,2} & 1 & \rho_{2,3} & \rho_{2,4} \\ \rho_{1,3} & \rho_{2,3} & 1 & \rho_{3,4} \\ \rho_{1,4} & \rho_{2,4} & \rho_{3,4} & 1 \end{bmatrix}$$

Examples of correlation structures

Compound symmetry (CS)

- constant correlation

$$\begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

Toeplitz (TOEP)

- correlation dependent on elapsed time

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix}$$

Unstructured (UN)

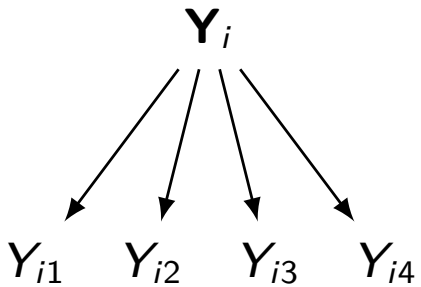
- flexible correlation over time

$$\begin{bmatrix} 1 & \rho_{1,2} & \rho_{1,3} & \rho_{1,4} \\ \rho_{1,2} & 1 & \rho_{2,3} & \rho_{2,4} \\ \rho_{1,3} & \rho_{2,3} & 1 & \rho_{3,4} \\ \rho_{1,4} & \rho_{2,4} & \rho_{3,4} & 1 \end{bmatrix}$$

Plan for today

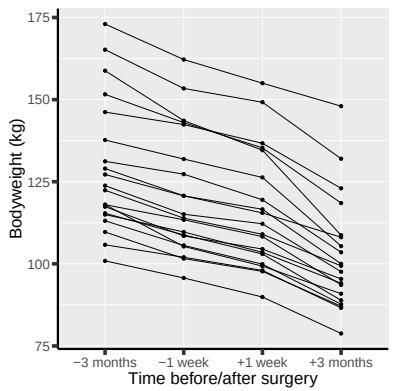
Graphical representation of correlation structures

- multilevel data
- nested vs. crossed correlation factors



Alternative specification of the covariance structure

- user friendly via a formula interface!
- necessary in some observational studies
  e.g. relate indirectly observed quantities
- complement the covariance patterns
  (CS, UN, . . .)

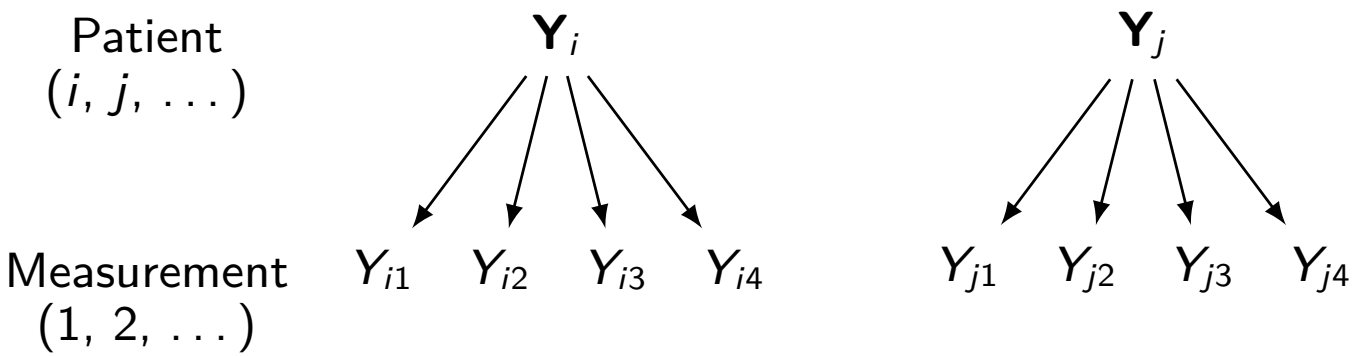


Hierarchical data

- graphical representation
- relation to correlation patterns

A graphical representation (2 levels)

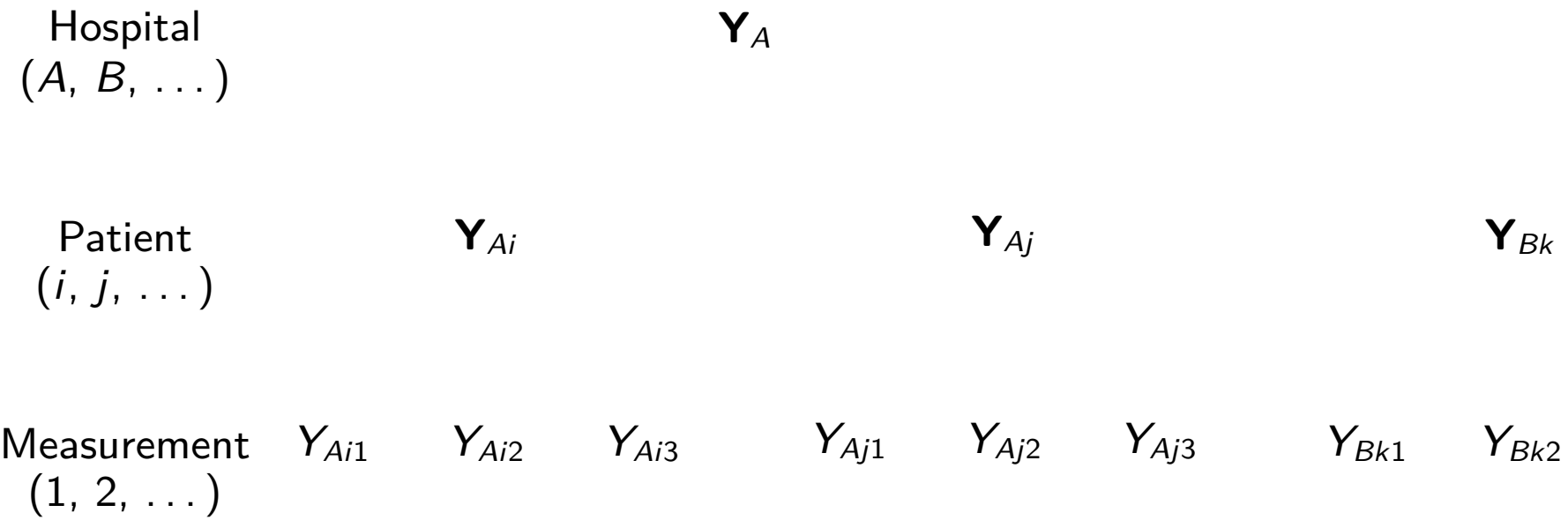
The **correlation** structure in the gastricbypass study can be viewed:



- **2 levels**: patient (level 1), measurement (level 0)
  $\rightarrow$ or family and individuals, doctor and patients, . . .
- observations are **correlated within cluster** (level 1)

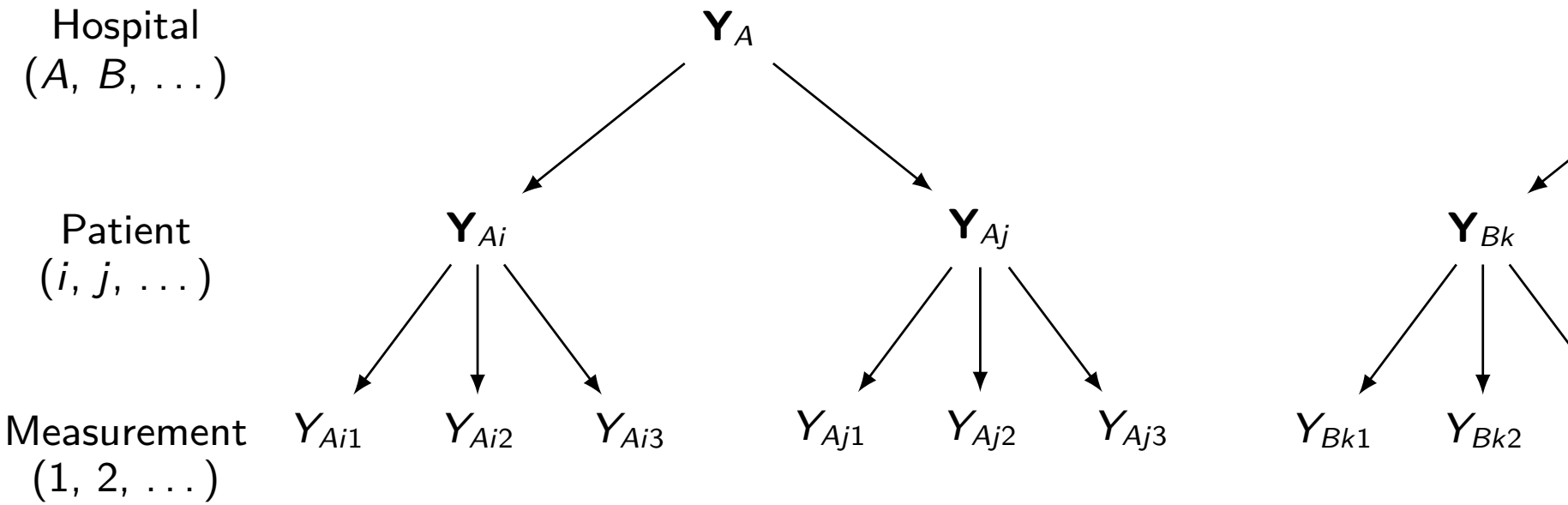
A graphical representation (3 levels, **nested**)

This graphical representation can be used with than 2 factors:



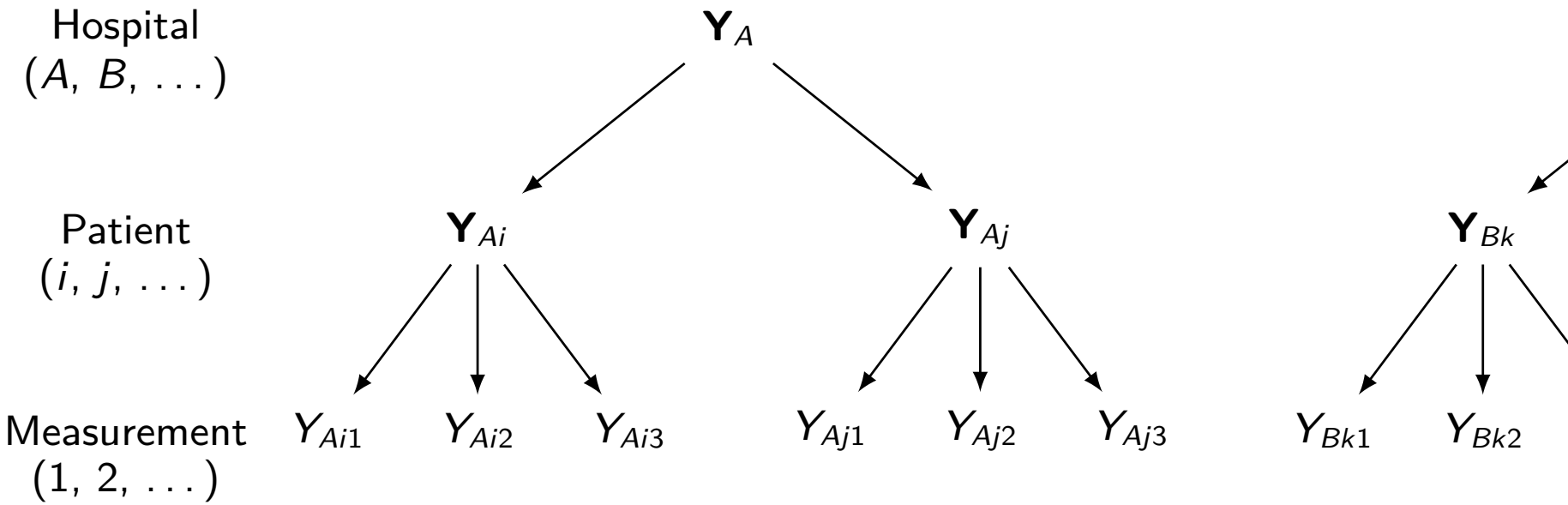
A graphical representation (3 levels, **nested**)

This graphical representation can be used with than 2 factors:



A graphical representation (3 levels, **nested**)

This graphical representation can be used with than 2 factors:

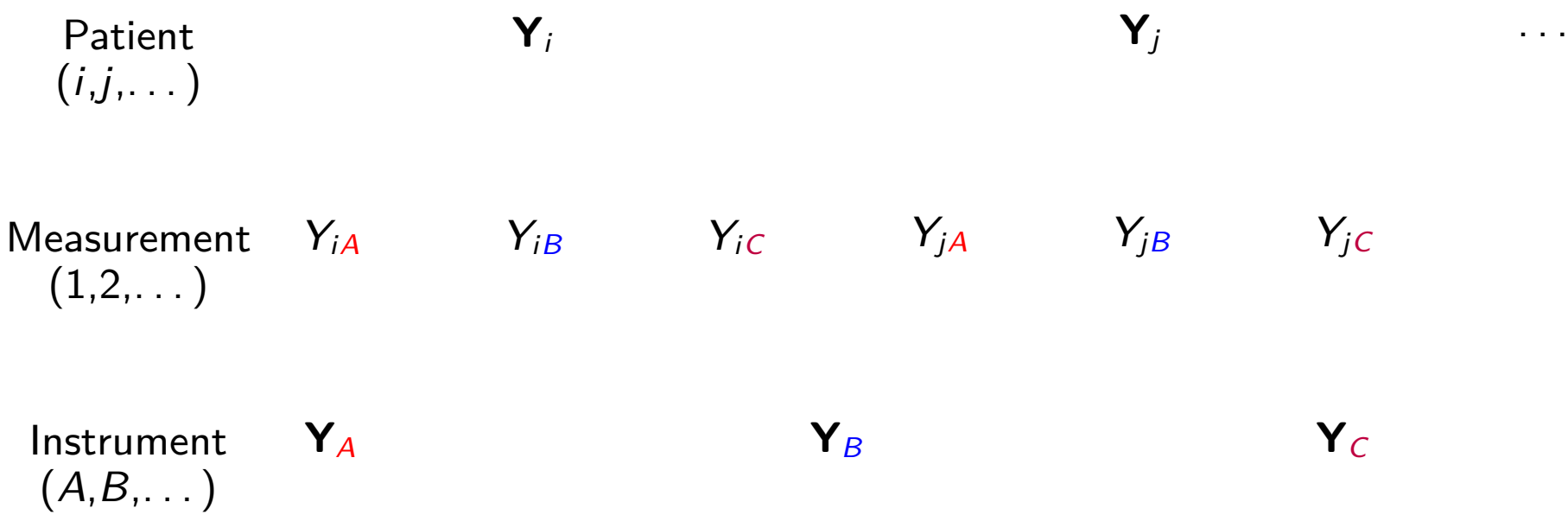


- **3 levels**: hospital (level 2), patient (level 1), measurement (level 0)
  $\rightarrow$ or school, class, student, . . .
- observation are correlated **within cluster**, and are
  even more correlated when belonging to the **same sub-cluster**

A graphical representation (3 levels, **crossed**)

This graphical representation does not require factors to be nested.

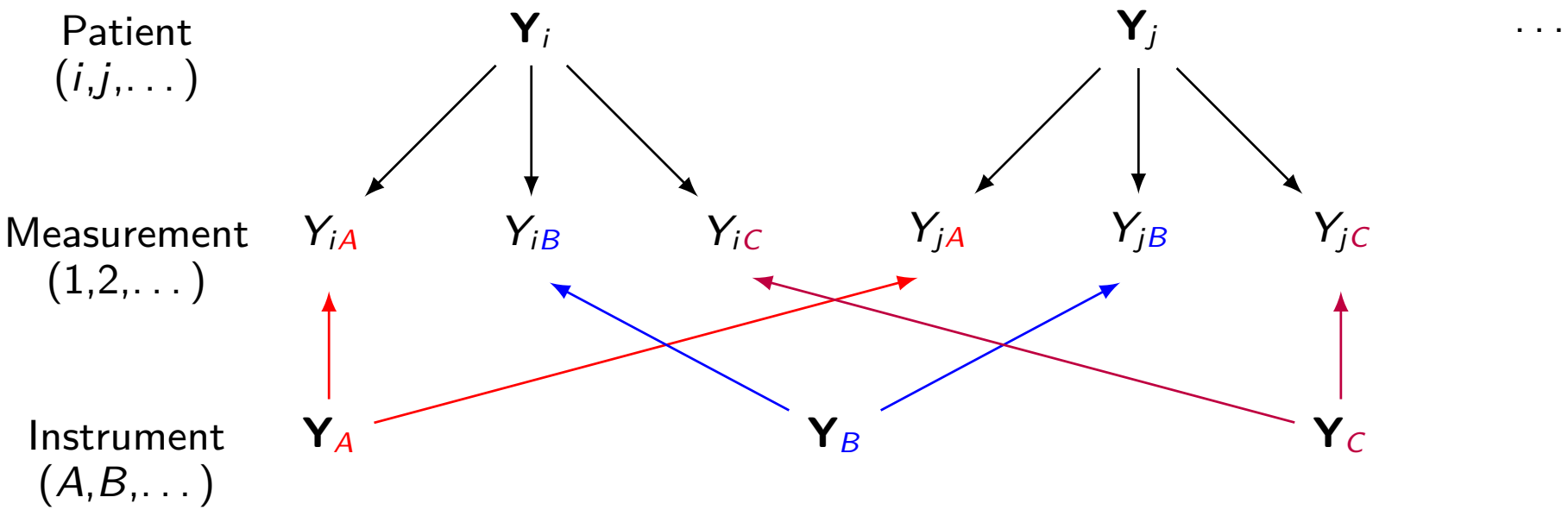
- different instruments are used during the study
- measurements are analyzed by batch in the lab'



A graphical representation (3 levels, **crossed**)

This graphical representation does not require factors to be nested.

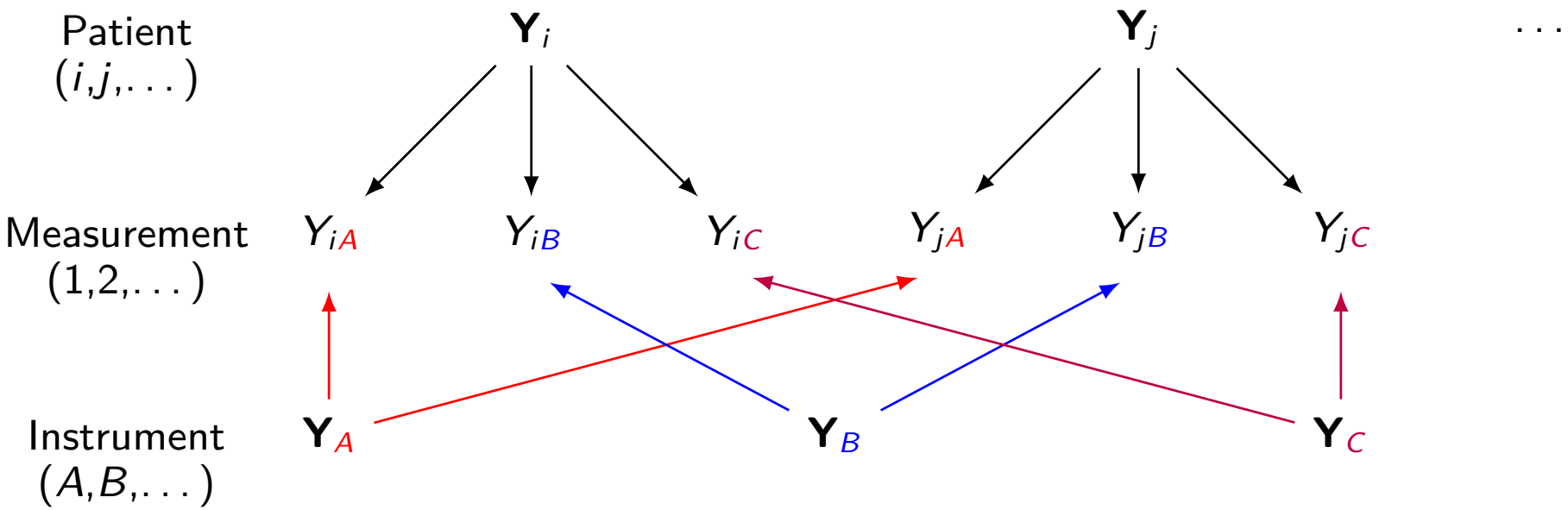
- different instruments are used during the study
- measurements are analyzed by batch in the lab'



A graphical representation (3 levels, **crossed**)

This graphical representation does not require factors to be nested.

- different instruments are used during the study
- measurements are analyzed by batch in the lab'



$\rightarrow$ observations are **correlated** if they originate from the same patient **or** made with the same instrument.

Note

Variables included in the graphical representation are special:

in comparison comparison to traditional covariates like sex:

Note

Variables included in the graphical representation are special:

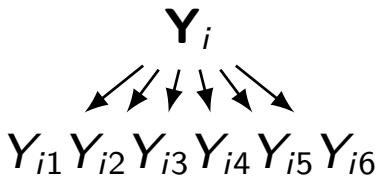
- **non-informative values**
  (patient id could be shuffled without consequence)
  (comparison patient 1 vs. 2 is not of interest)
- **range growing with the sample size**
  (more data implies more patients id's)

in comparison comparison to traditional covariates like sex:

- admit only two values (male and female),
  independently of the sample size.
- each has a specific meaning which may be interest, e.g.
  compare weight loss for male vs. female.

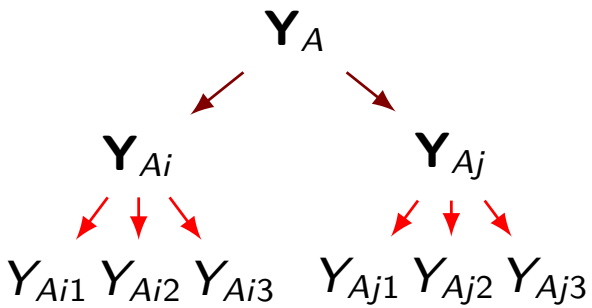
Who is what 

2 level



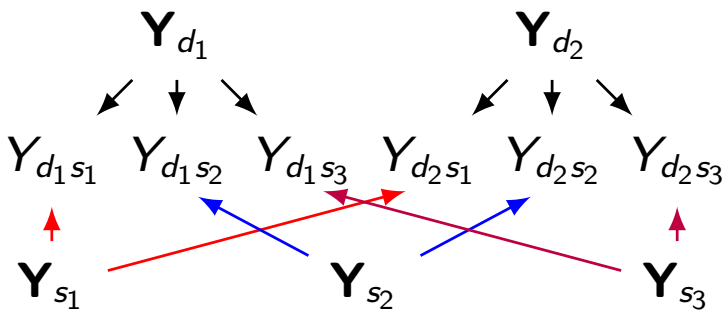
$$R_A = \begin{bmatrix} 1 & \rho_d & \rho_d & \rho_s & 0 & 0 \\ \rho_d & 1 & \rho_d & 0 & \rho_s & 0 \\ \rho_d & \rho_d & 1 & 0 & 0 & \rho_s \\ \rho_s & 0 & 0 & 1 & \rho_d & \rho_d \\ 0 & \rho_s & 0 & \rho_d & 1 & \rho_d \\ 0 & 0 & \rho_s & \rho_d & \rho_d & 1 \end{bmatrix}$$

3 level (nested)



$$R_B = \begin{bmatrix} 1 & \rho & \rho & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho & \rho & \rho \\ \rho & \rho & 1 & \rho & \rho & \rho \\ \rho & \rho & \rho & 1 & \rho & \rho \\ \rho & \rho & \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & \rho & \rho & 1 \end{bmatrix}$$

3 level (crossed)



$$R_C = \begin{bmatrix} 1 & \rho_p & \rho_p & \rho_h & \rho_h & \rho_h \\ \rho_p & 1 & \rho_p & \rho_h & \rho_h & \rho_h \\ \rho_p & \rho_p & 1 & \rho_h & \rho_h & \rho_h \\ \rho_h & \rho_h & \rho_h & 1 & \rho_p & \rho_p \\ \rho_h & \rho_h & \rho_h & \rho_p & 1 & \rho_p \\ \rho_h & \rho_h & \rho_h & \rho_p & \rho_p & 1 \end{bmatrix}$$

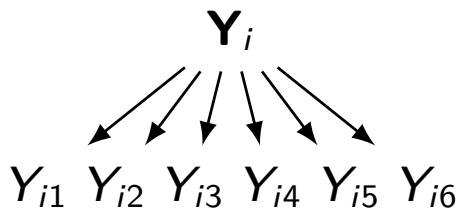
Correlation structure (2 levels)

⚠ Simplifying assumption:

- within factor(s), observations are all equally correlated

Patient

Measurement

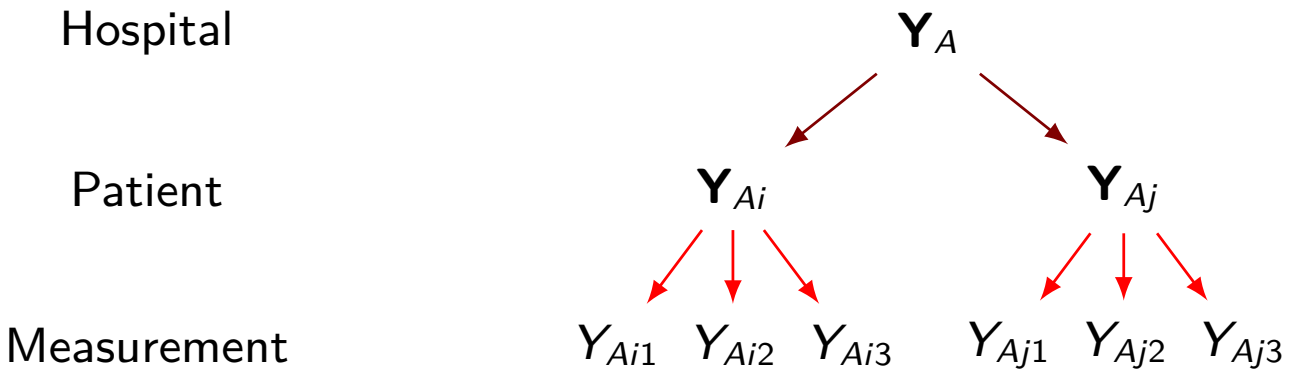


$$R = \begin{bmatrix} 1 & \rho & \rho & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho & \rho & \rho \\ \rho & \rho & 1 & \rho & \rho & \rho \\ \rho & \rho & \rho & 1 & \rho & \rho \\ \rho & \rho & \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & \rho & \rho & 1 \end{bmatrix}$$

Correlation structure (3 levels, **nested**)

⚠ Simplifying assumptions:

- within factor(s), observations are all equally correlated

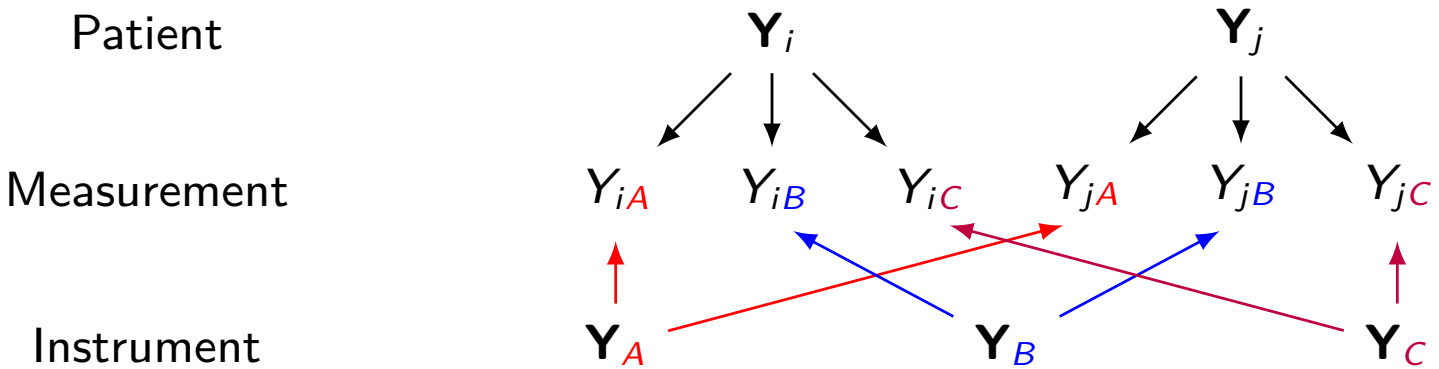


$$R = \begin{bmatrix} 1 & \rho_p & \rho_p & \rho_h & \rho_h & \rho_h \\ \rho_p & 1 & \rho_p & \rho_h & \rho_h & \rho_h \\ \rho_p & \rho_p & 1 & \rho_h & \rho_h & \rho_h \\ \rho_h & \rho_h & \rho_h & 1 & \rho_p & \rho_p \\ \rho_h & \rho_h & \rho_h & \rho_p & 1 & \rho_p \\ \rho_h & \rho_h & \rho_h & \rho_p & \rho_p & 1 \end{bmatrix}$$

Correlation structure (3 levels, **crossed**)

⚠ Simplifying assumption:

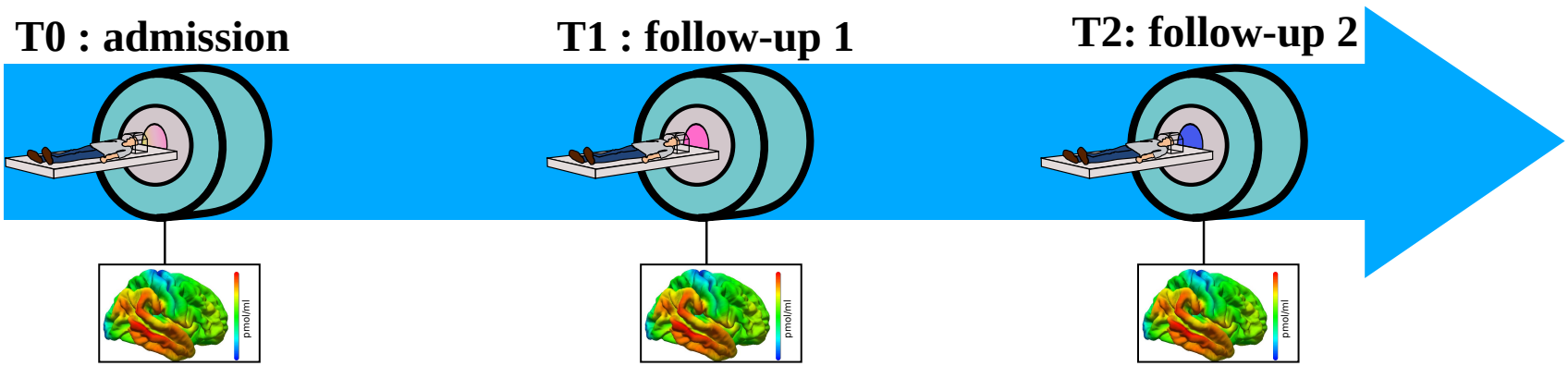
- within factor(s), observations are all equally correlated



$$R = \begin{bmatrix} 1 & \rho_d & \rho_d & \rho_s & 0 & 0 \\ \rho_d & 1 & \rho_d & 0 & \rho_s & 0 \\ \rho_d & \rho_d & 1 & 0 & 0 & \rho_s \\ \rho_s & 0 & 0 & 1 & \rho_d & \rho_d \\ 0 & \rho_s & 0 & \rho_d & 1 & \rho_d \\ 0 & 0 & \rho_s & \rho_d & \rho_d & 1 \end{bmatrix}$$

Exercise & break!

Consider a study where we would measure some brain signal repeatedly over time:



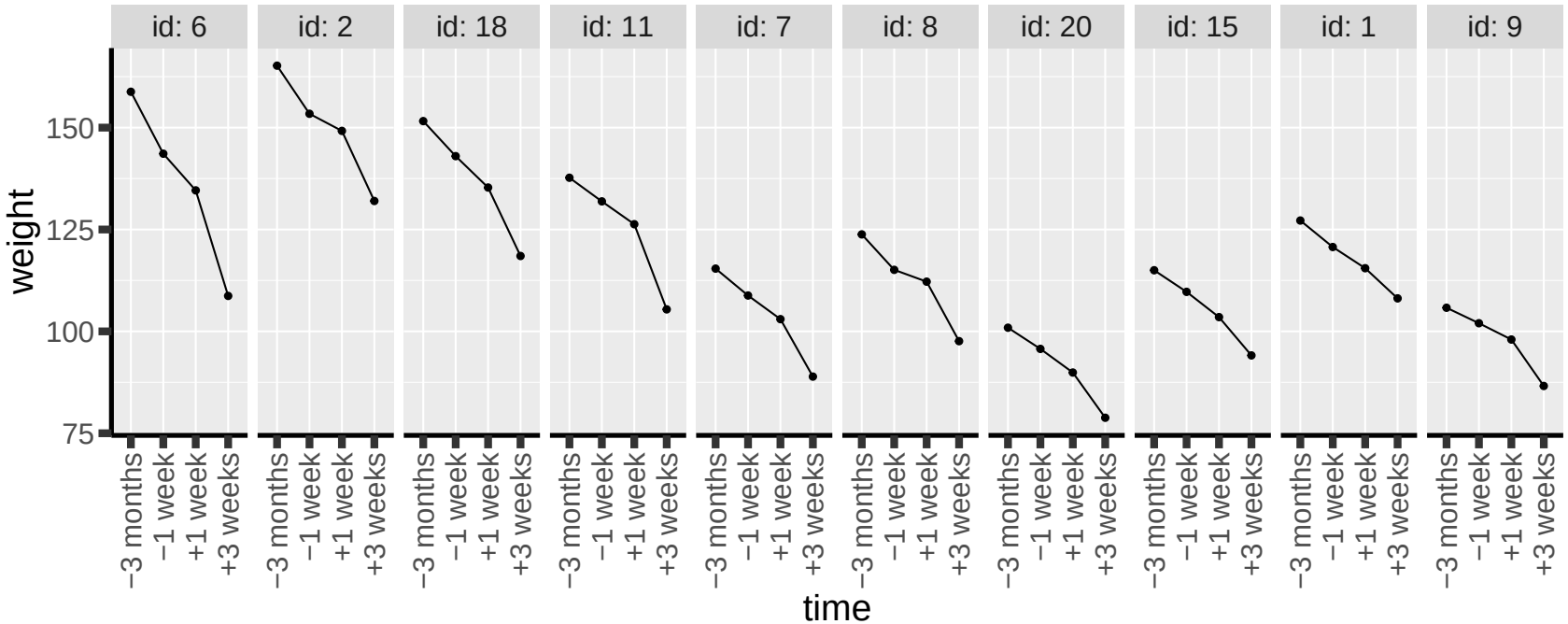
How would you represent the correlation structure 

- considering a single brain region
- considering a multiple brain regions

Latent variables

- as a part of the research question
    - as a way to account for correlation

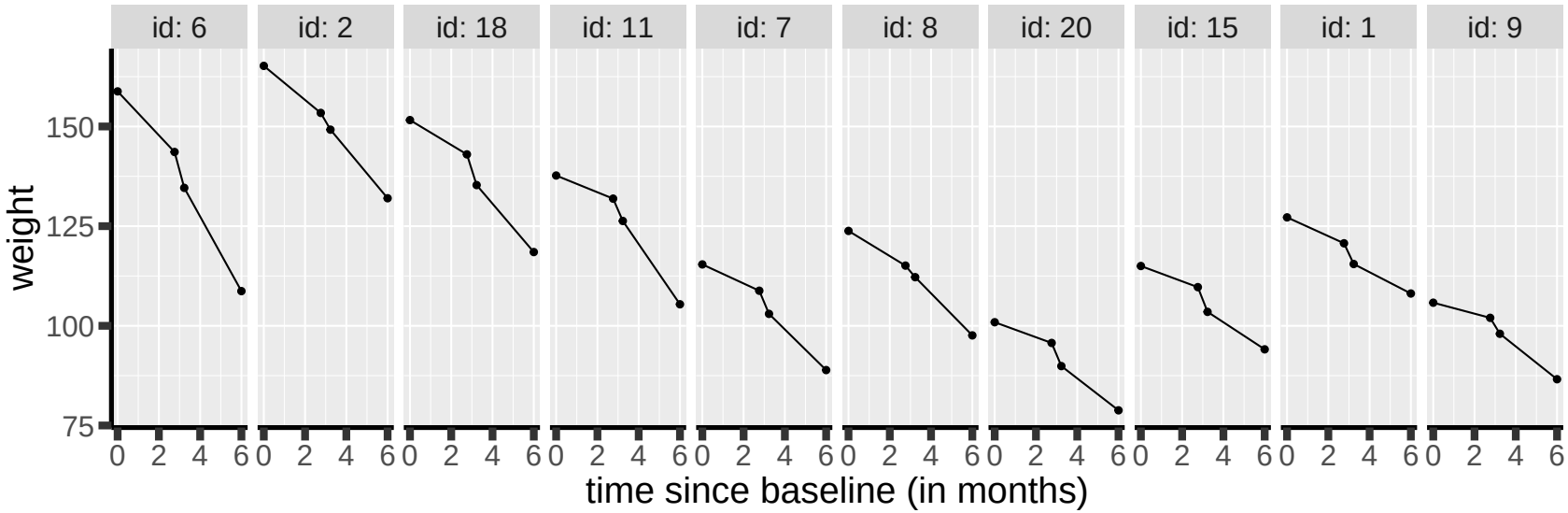
An unusual research question



Research question: is weight loss associated with the initial weight?

- weight loss is not directly measured but can be deduced from the observed weight over time

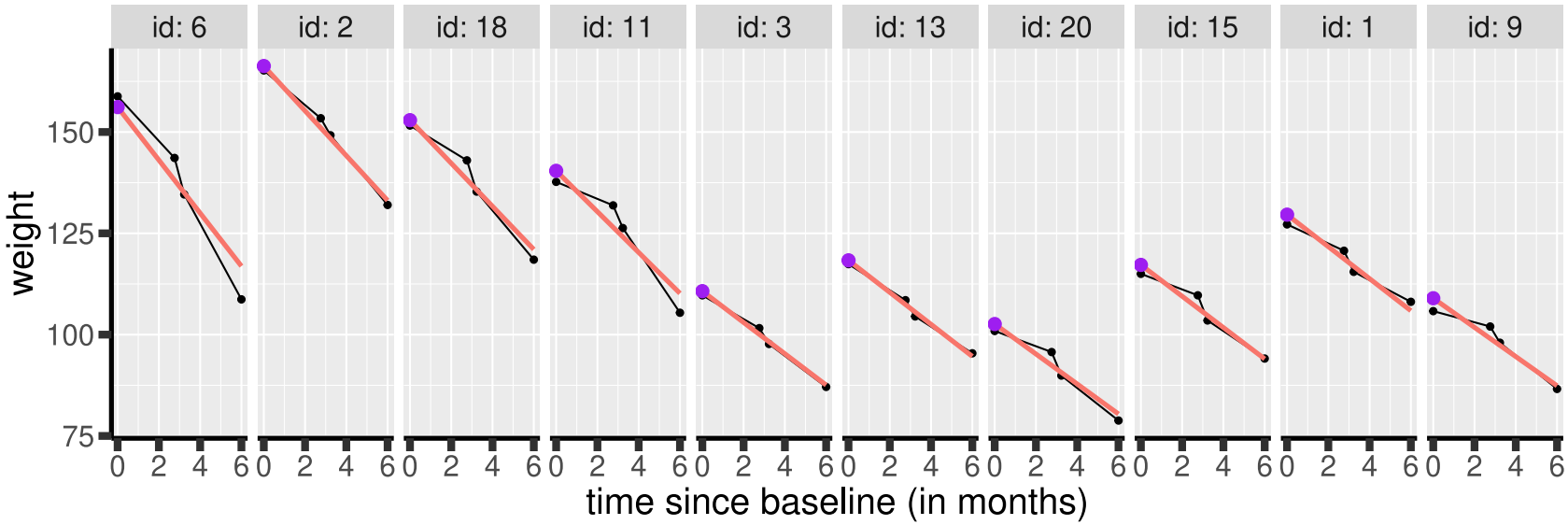
An unusual research question



Research question: is weight loss associated with the initial weight?

- weight loss is not directly measured but can be deduced from the observed weight over time

An unusual research question



Research question: is weight loss associated with the initial weight?

- weight loss is not directly measured but can be deduced from the observed weight over time

$\rightarrow$ slope associated with the intercept?

A two-step approach

Run a linear regression for each individual:

- intercept and time as a continuous variable

```
  id (Intercept)  time.num
1  1    127.5615 -3.231514
2  2    166.6006 -5.554824
3  3    110.3994 -3.794641
```

Compute the Pearson correlation between both quantities:

```
[1] -0.6217227
```

A two-step approach

Run a linear regression for each individual:

- intercept and time as a continuous variable

```
  id (Intercept)  time.num
1  1    127.5615 -3.231514
2  2    166.6006 -5.554824
3  3    110.3994 -3.794641
```

Compute the Pearson correlation between both quantities:

```
[1] -0.6217227
```

⚠ **p-values/CI from usual tests should not be trusted!**
Indeed modeling uncertainty about the intercept and slope is not accounted for

Single-step approach: random slope model

In SAS:

```
proc mixed data=gastricbypassL;
class id time;
model weight = time.num / ddfm=kr cl;
random intercept time.num /
       type=un subject=id g gcor v vcorr;
  run;
```

In R

```
e.rs <- lmer(weight ~ time.num + (time.num|id),
      data = gastricbypassL)
```

Latent variable in a random slope model

The model estimate a population intercept and slope:

```
fixef(e.rs)
```

```
(Intercept)    time.num
 130.496885   -4.479778
```

as well as a subject specific intercept and slope:

```
coef(e.rs)
```

```
  (Intercept)  time.num
1    129.6055 -3.949467
2    166.2820 -5.530303
3    110.7275 -3.861765
```

Example of software output (in R)

Estimate correlation:

```
VarCorr(e.rs)
```

```
Groups   Name        Std.Dev. Corr
id       (Intercept) 20.21900
         time.num     0.88028 -0.731
Residual              3.20841
```

with its confidence interval:

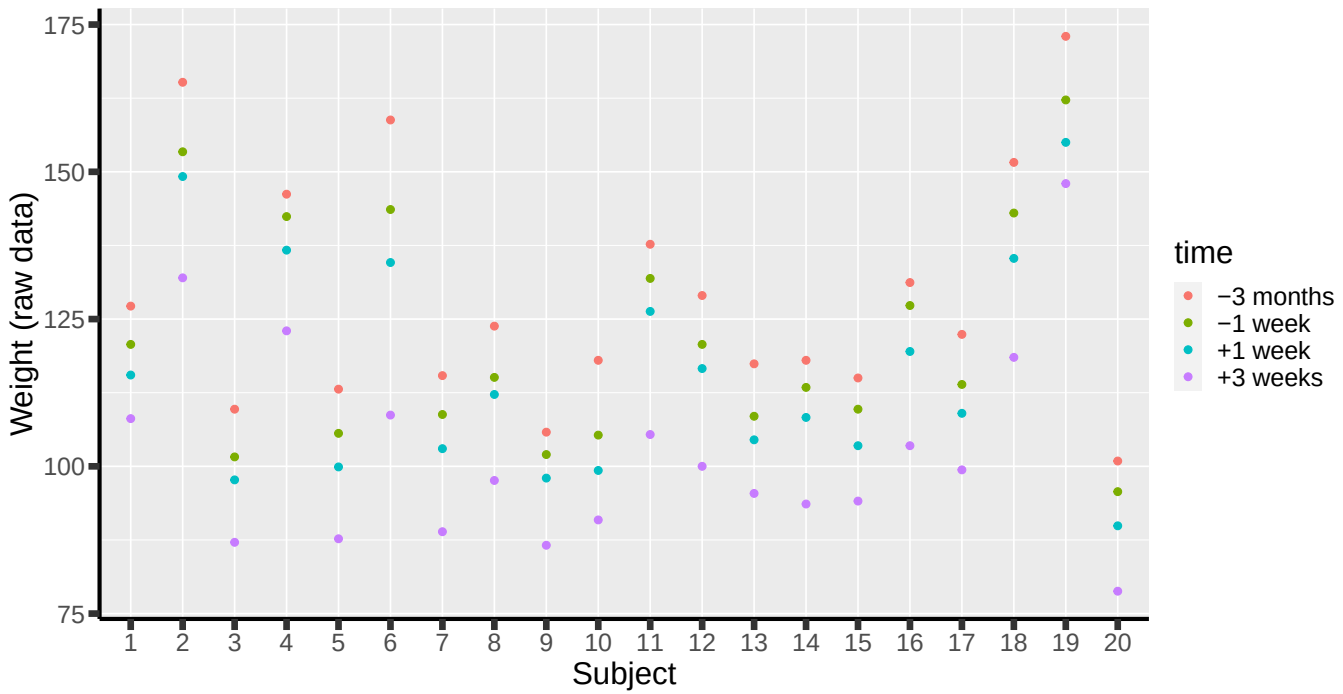
```
confint(e.rs, oldNames=FALSE)[2,,drop=FALSE]
```

```
Computing profile confidence intervals ...
                                2.5 %    97.5 %
cor_time.num.(Intercept)|id    -1 -0.268034
Warning messages:
1: In FUN(X[[i]], ...) :
  non-monotonic profile for cor_time.num.(Intercept)|id
2: In confint.thpr(pp, level = level, zeta = zeta) :
  bad spline fit for cor_time.num.(Intercept)|id: falling back t
```

Intuition about random effects

Two viewpoints:

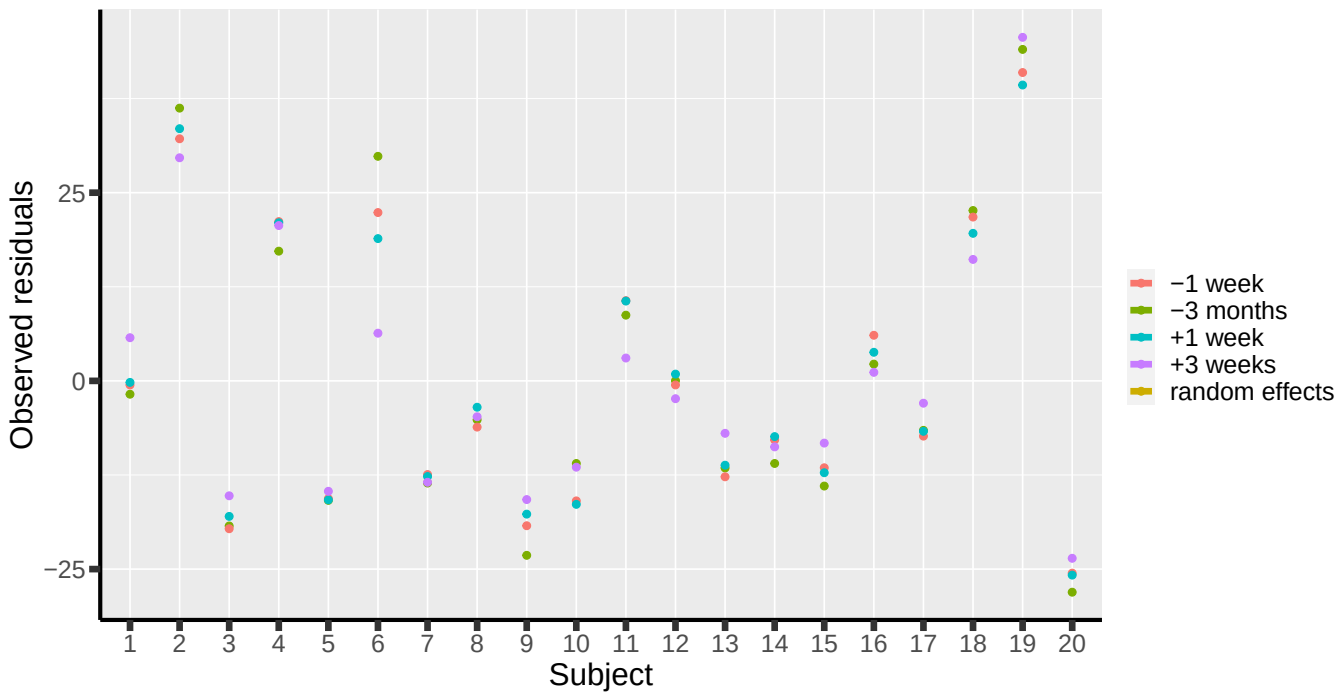
- weight measurements are correlated within patient
- each patient has his own "natural" weight level/slope



Intuition about random effects

Two viewpoints:

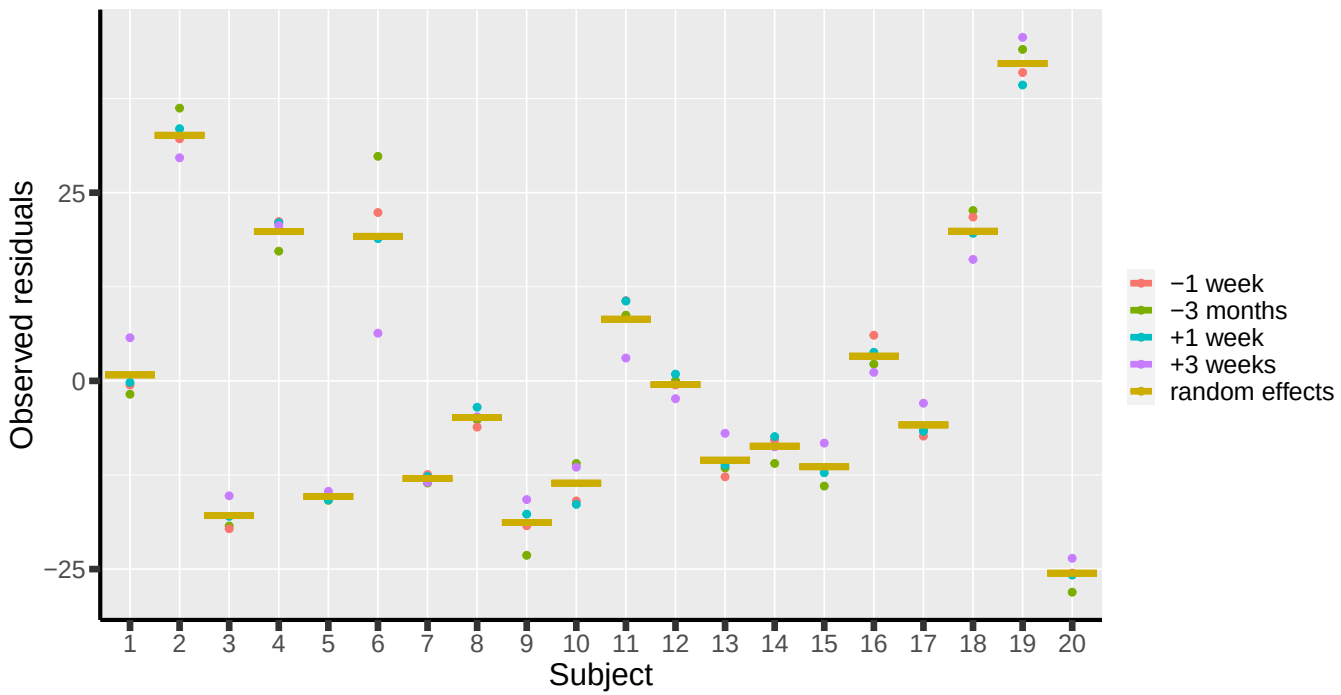
- weight measurements are correlated within patient
- each patient has his own "natural" weight level/slope



Intuition about random effects

Two viewpoints:

- weight measurements are correlated within patient
- each patient has his own "natural" weight level/slope

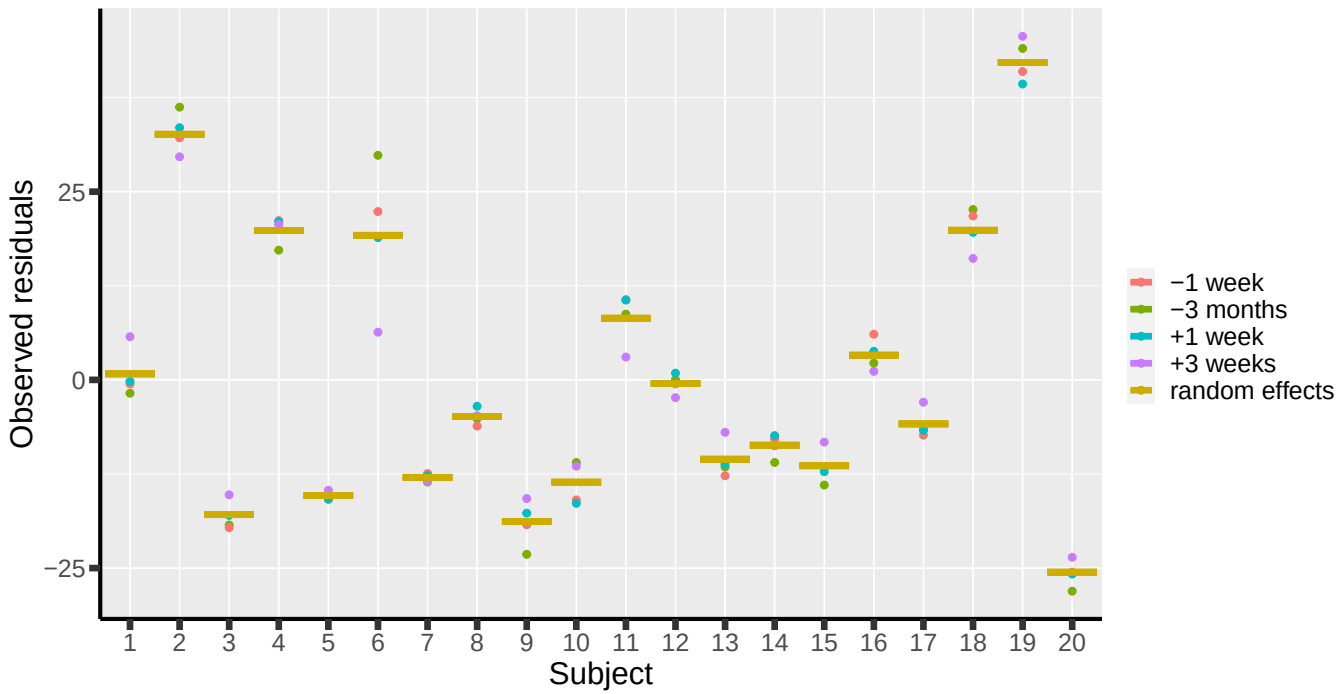


Random effect: systematic deviation from the population weight

Intuition about random effects

Two viewpoints:

- weight measurements are correlated within patient
  observed residuals $\varepsilon_{it} = u_i + \xi_{it}$ are (very) correlated!
- each patient has his own "natural" weight level/slope
  conditional residuals ξ_{it} are white noise



Random effect: systematic deviation from the population weight

Random effect models

- random intercept model
- random slope model

Random intercept model

For individual i at time t, we model the weight Y_{it} as:

$$Y_{it} = \mu_t + u_i + \xi_{it}$$

where:

- μ_t is the population mean over time (fixed effect).
- $u_i \sim \mathcal{N}(0, \sigma_B^2)$ is the random intercept for subject i
  $\rightarrow$ mathematical construct which may reflect unobservable concepts (e.g. "natural" weight)
- $\boldsymbol{\xi}_i = (\xi_{i1}, \xi_{i2}, \xi_{i3}, \xi_{i4}) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_W^2 & 0 & 0 & 0 \\ 0 & \sigma_W^2 & 0 & 0 \\ 0 & 0 & \sigma_W^2 & 0 \\ 0 & 0 & 0 & \sigma_W^2 \end{bmatrix} \right)$
  residual error, independent of u_i

Random intercept model (software)

In R :

```
library(lme4)
library(lmerTest) ## P-value based on KR approximation
e.ReI <- lmer(weight ~ time + (1|id),
       data = gastricbypassL)
```

In SAS:

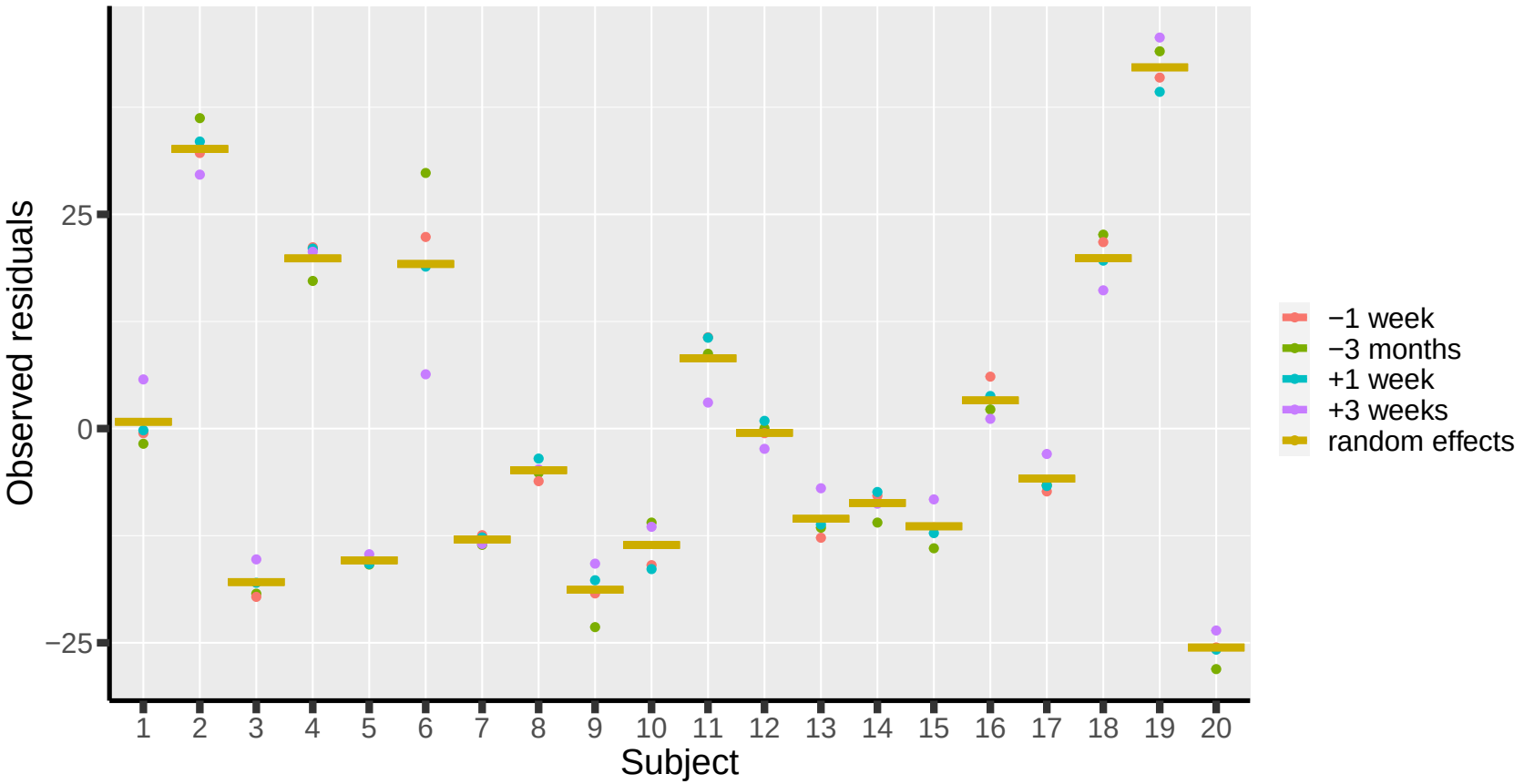
```
proc mixed data=gastricbypassL;
class id time;
model weight = time / ddfm=kr cl;
random id;
run;
```

Random intercept model (output)

```
e.ReI
```

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: weight ~ time + (1 | id)
   Data: gastricbypassL
REML criterion at convergence: 499.0104
Random effects:
 Groups   Name        Std.Dev.
 id       (Intercept) 18.379
 Residual              3.246
Number of obs: 80, groups:  id, 20
Fixed Effects:
 (Intercept)   time-1 week   time+1 week  time+3 weeks
      128.97         -7.73        -13.27        -26.60
```

Variance decomposition (visually)



- σ_B^2: variance of the random effect (i.e. yellow lines)
- σ_W^2: variance of the noise (i.e. points around the yellow lines)

Variance decomposition (numerically)

Variation	Variance component	Estimate	Proportion of variance
Between	σ_B^2	$18.379^2 \approx 337$	96.96%
Within	σ_W^2	$3.246^2 \approx 11$	3.02%
Total	$\sigma_B^2 + \sigma_W^2$	$18.379^2 + 3.246^2 \approx 348$	100%

[1] at the same time or once removing the time effect

Variance decomposition (numerically)

Variation	Variance component	Estimate	Proportion of variance
Between	σ_B^2	$18.379^2 \approx 337$	96.96%
Within	σ_W^2	$3.246^2 \approx 11$	3.02%
Total	$\sigma_B^2 + \sigma_W^2$	$18.379^2 + 3.246^2 \approx 348$	100%

Typical difference in weight between two measurements[1]:

- on the **same** patient:

- between **two** different patients:

[1] at the same time or once removing the time effect

Variance decomposition (numerically)

Variation	Variance component	Estimate	Proportion of variance
Between	σ_B^2	$18.379^2 \approx 337$	96.96%
Within	σ_W^2	$3.246^2 \approx 11$	3.02%
Total	$\sigma_B^2 + \sigma_W^2$	$18.379^2 + 3.246^2 \approx 348$	100%

Typical difference in weight between two measurements[1]:

- on the **same** patient:

$$\pm 1.96\sqrt{2\sigma_W^2} \approx \pm 9.19$$

- between **two** different patients:

$$\pm 1.96\sqrt{2(\sigma_B^2 + \sigma_W^2)} \pm 51.71$$

[1] at the same time or once removing the time effect

Intra-class correlation (ICC)

We would like to evaluate the correlation between two observations from the same patient (at different time t and t')

We can compute:

- the variance $\mathbb{V}ar\left[Y_{it}\right] = \mathbb{V}ar\left[Y_{it'}\right] =$
- the covariance

$$\mathbb{C}ov\left(Y_{it}, Y_{it'}\right) =$$

So the correlation is

$$\rho = \mathbb{C}or\left(Y_{it}, Y_{it'}\right) =$$

Intra-class correlation (ICC)

We would like to evaluate the correlation between two observations from the same patient (at different time t and t')

We can compute:

- the variance $\mathbb{V}ar\left[Y_{it}\right] = \mathbb{V}ar\left[Y_{it'}\right] = \sigma_B^2 + \sigma_W^2$
- the covariance

$$\begin{aligned}\mathbb{C}ov\left(Y_{it}, Y_{it'}\right) =& \mathbb{C}ov\left(\mu_t + u_i + \xi_{it}, \mu_{t'} + u_i + \xi_{it'}\right) \\ =& \mathbb{C}ov\left(u_i, u_i\right) = \sigma_B^2\end{aligned}$$

So the correlation is

$$\rho = \mathbb{C}or\left(Y_{it}, Y_{it'}\right) = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} = 0.970$$

Relation to covariance pattern (in formula)

We have seen that:

- the variance is $\sigma_B^2 + \sigma_W^2$
- the covariance is σ_B^2

So:

$$\Omega = \begin{bmatrix} \sigma_W^2 + \sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\ \sigma_B^2 & \sigma_W^2 + \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\ \sigma_B^2 & \sigma_B^2 & \sigma_W^2 + \sigma_B^2 & \sigma_B^2 \\ \sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_W^2 + \sigma_B^2 \end{bmatrix}$$

Relation to covariance pattern (in formula)

We have seen that:

- the variance is $\sigma_B^2 + \sigma_W^2$
- the covariance is σ_B^2

So:

$$\Omega = \begin{bmatrix} \sigma_W^2 + \sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\ \sigma_B^2 & \sigma_W^2 + \sigma_B^2 & \sigma_B^2 & \sigma_B^2 \\ \sigma_B^2 & \sigma_B^2 & \sigma_W^2 + \sigma_B^2 & \sigma_B^2 \\ \sigma_B^2 & \sigma_B^2 & \sigma_B^2 & \sigma_W^2 + \sigma_B^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

with $\sigma^2 = \sigma_W^2 + \sigma_B^2$.

Exaclty a compound symmetry structure!
(assuming positive correlation)

Relation to covariance pattern (in R)

```
e.CS <- lmm(weight ~ time,
     repetition =~time|id,
     structure = "CS",
     data = gastricbypassL)
logLik(e.CS)
logLik(e.lme)
```

```
[1] -249.5052
'log Lik.' -249.5052 (df=6)
```

```
coef(e.CS, effects = "all", transform.sigma="square")
```

```
(Intercept)   time-1 week  time+ 1 week time+ 3 weeks
128.9700000    -7.7300000   -13.2700000   -26.6050000
    sigma^2           rho
348.3162479     0.9697449
```

Marginal vs. conditional predictions

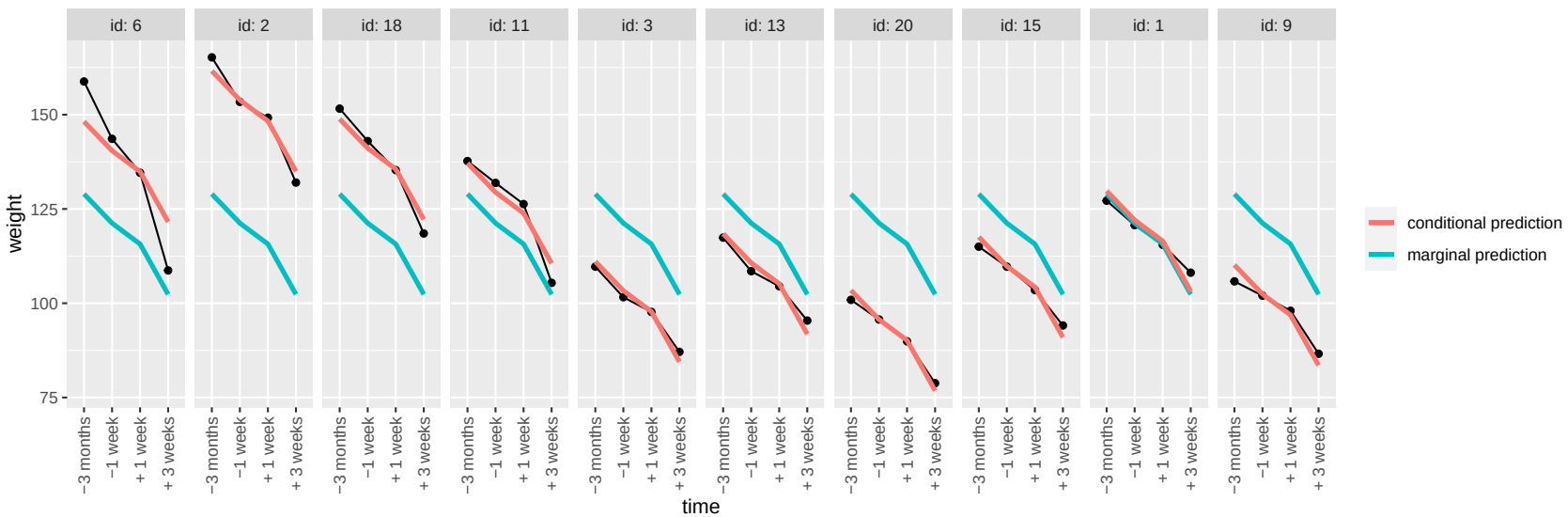
When using mixed model, one can predict:

- for a new individual: $\widehat{Y}_{it} = \widehat{\mu}_t$

```
predict(e.ReI, newdata = gastricbypassL, re.form=~0)
```

- for a known individual (BLUP): $\widehat{Y}_{it} = \widehat{\mu}_t + \widehat{u}_i$

```
predict(e.ReI, newdata = gastricbypassL)
```



What are the random effects?

In this simple model,

$$u_i = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2/4} \times \frac{1}{4} \sum_{t=1}^{4} (Y_{it} - \mu_t)$$

i.e. 0.992 times the average residual for individual i.
$\rightarrow$ average of the residuals shrinked toward 0
$\rightarrow$ individual predictions are shrinked toward the population mean
σ_B^2 controls the shrinkage/how individual-specific predictions are

Numerically:

```
 [1]   0.8  32.6 -17.9  19.9 -15.4  19.2 -12.9  -4.9 -18.8 -13.6
[11]   8.2  -0.5 -10.5  -8.7 -11.4   3.3  -5.8  19.9  42.2 -25.5
```

Break & exercise

Discuss with your neighbor:

- benefits and limitations of the random intercept model
- why does a random intercept model fail to be estimated when considering the outcome glucagon?
  (pre- vs. post. operation)

```
df <- gastricbypassL[gastricbypassL$visit %in% 2:3,]
lmer(glucagonAUC ~ time + (1|id), data = df)
```

```
boundary (singular) fit: see help('isSingular')
```

Limitations of random intercept models

A random intercept model assumes:

- **equal, positive correlation** between any two replicates
  $\rightarrow$ unrealistic: close in time means larger correlation
- **same variance** for all replicates
  $\rightarrow$ unrealistic: variability changes over time

Limitations of random intercept models

A random intercept model assumes:

- **equal, positive correlation** between any two replicates
  $\rightarrow$ unrealistic: close in time means larger correlation
- **same variance** for all replicates
  $\rightarrow$ unrealistic: variability changes over time

Can be improved by adding more random effects:

- including individual-specific time changes

Random slope models (model)

For individual i at time t we model its weight Y_{it} as:

$$Y_{it} = \mu_t + u_i + t \times v_i + \xi_{it}$$

where:

- $u_i \sim \mathcal{N}(0, \sigma_I^2)$: random intercept for subject i
- $v_i \sim \mathcal{N}(0, \sigma_S^2)$: random slope for subject i
- $\xi_{it} \sim \mathcal{N}(0, \sigma_W^2)$: residual error term for subject i at time t

Random slope models (model)

For individual i at time t we model its weight Y_{it} as:

$$Y_{it} = \mu_t + u_i + t \times v_i + \xi_{it}$$

where:

- $u_i \sim \mathcal{N}(0, \sigma_I^2)$: random intercept for subject i
- $v_i \sim \mathcal{N}(0, \sigma_S^2)$: random slope for subject i
- $\xi_{it} \sim \mathcal{N}(0, \sigma_W^2)$: residual error term for subject i at time t

Usually we model the correlation between u_i and v_i (parameter ρ)

- individuals with high baseline weight tend to lose more weight

⚠ Not to be confused with the ρ parameters in the correlation structure!

Random slope models (software)

```
## time.num 0    : -3 months,    2.76: -1 week
##          3.23: +1 week,      6    : +3 months
```

in R :

```
e.ReS <- lmer(weight ~ time + (time.num|id),
       data = gastricbypassL)
```

in SAS:

```
proc mixed data=gastricbypassL;
class id time;
model weight = time / ddfm=kr cl;
random intercept time.num /
       type=un subject=id g gcor v vcorr;
run;
```

Random slope models (output)

```
e.ReS
```

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: weight ~ time + (time.num | id)
   Data: gastricbypassL
REML criterion at convergence: 469.0984
Random effects:
 Groups   Name        Std.Dev. Corr
 id       (Intercept) 20.342
          time.num     1.069   -0.65
 Residual              1.909
Number of obs: 80, groups:  id, 20
Fixed Effects:
 (Intercept)   time-1 week   time+1 week  time+3 weeks
      128.97         -7.73        -13.27        -26.60
```

Comments

Compared to the random intercept model:

- the fixed effect are almost identical
- the residual variance is even smaller
  (1.909^2 instead of 3.246^2)
- the variance of the random intercept is slightly larger
  (20.342^2 instead of 18.379^2)
  $\rightarrow$ more flexibility helps to capture individual variations

Comments

Compared to the random intercept model:

- the fixed effect are almost identical
- the residual variance is even smaller
  (1.909^2 instead of 3.246^2)
- the variance of the random intercept is slightly larger
  (20.342^2 instead of 18.379^2)
  $\rightarrow$ more flexibility helps to capture individual variations

⚠ the variance of the random slope is not comparable to the other variance as it is **by time unit**
(here 1.069^2 per month).
$\rightarrow$ difference in slope between individuals
$\pm 1.96\sqrt{2\sigma_S^2} \approx \pm 2.96$ kg/month

Correlation between random effects

```
VarCorr(e.ReS)
```

```
Groups   Name        Std.Dev. Corr
id       (Intercept) 20.3417
         time.num     1.0687  -0.649
Residual              1.9089
```

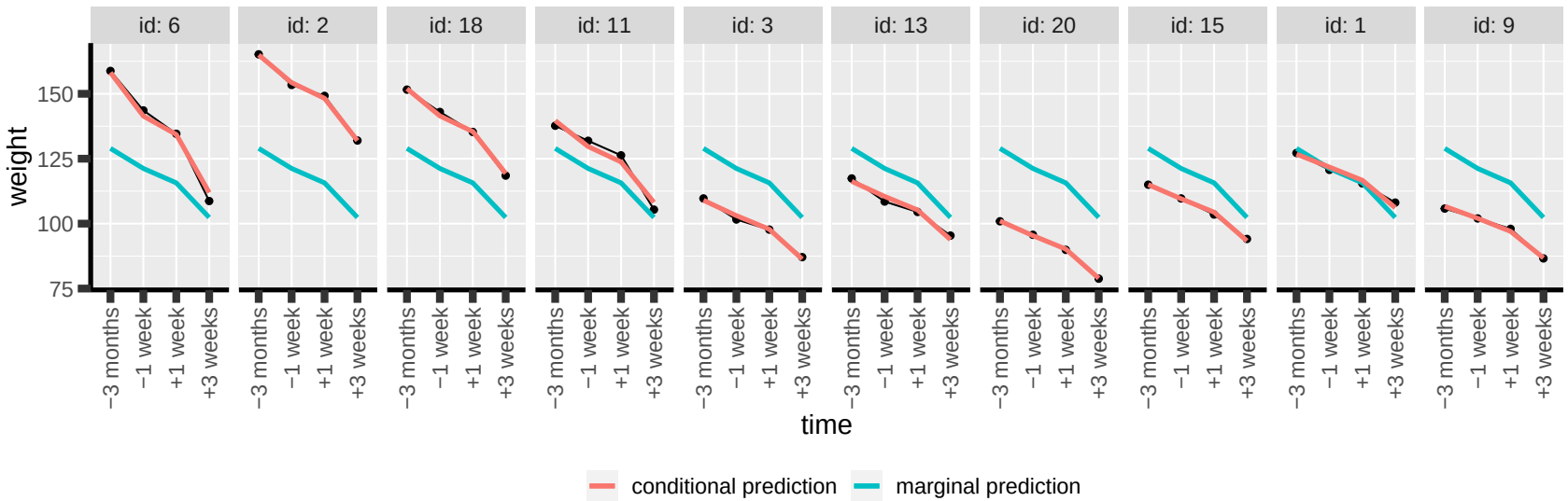
```
confint(e.ReS, oldNames=FALSE)[2,,drop=FALSE]
```

```
Computing profile confidence intervals ...
                                    2.5 %     97.5 %
cor_time.num.(Intercept)|id -0.8555569 -0.2712663
```

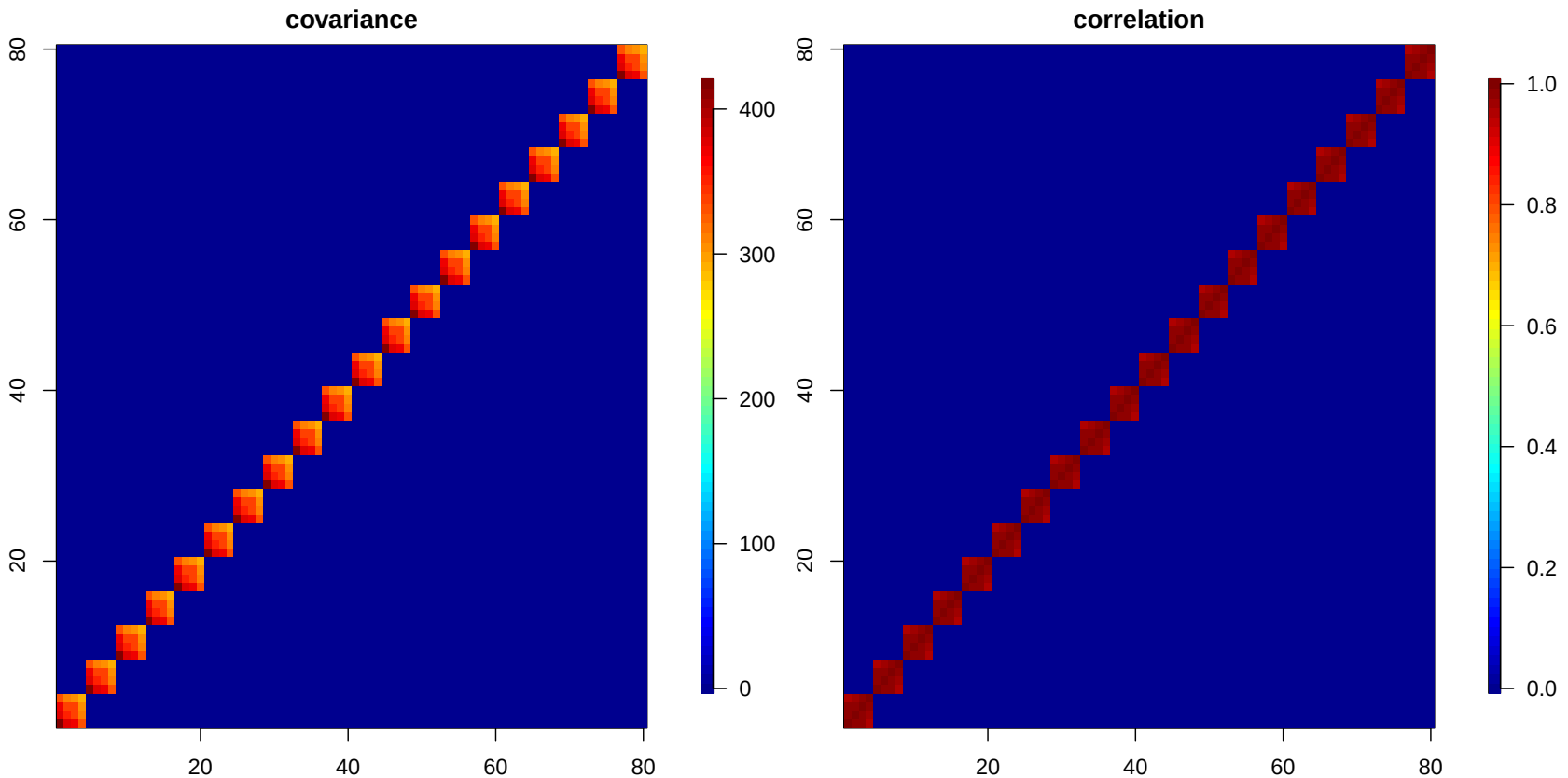
Evidence for a correlation between the random intercept and random slope

- larger initial weight is associated with larger weight loss

Random slope models (illustration)



What about Ω and R?



What about Ω and R?

The residual variance-covariance matrix has a complex expression:

- depends on 4 parameters (σ_I^2, σ_S^2, ρ, σ_W^2)
- more flexible than CS but less than UN

```
correlation                                                  standard
            -3 months   -1 week   +1 week  +3 weeks     deviation
-3 months 1.0000000 0.9831382 0.9799849 0.9472105      20.43121
-1 week   0.9831382 1.0000000 0.9891228 0.9739420      18.66236
+1 week   0.9799849 0.9891228 1.0000000 0.9773993      18.39135
+3 weeks  0.9472105 0.9739420 0.9773993 1.0000000      17.00906
```

Exercise: numerical vs. categorical variable

In the gastricbypass study we can consider time:

- as a numeric variable `time.num`
- as a categorical variable `time`

What is the difference between following mixed models:

```
weight ~ time.num + (time.num|id)
weight ~ time + (time.num|id)
weight ~ time.num + (time|id)
weight ~ time + (time|id)
```

Hint: only two of them make sense!

Conclusion

When using a mixed model, one has to specify:

- a model for the mean "same as usual"
- a model for the variance and correlation "many possibilities!"

Direct specification:

- Unstructured pattern (UN) is a good default for discrete time/balanced design

Indirect specification via random effects:

- random slopes models are well suited for time series
- nested/crossed random effects for complex hierarchical structure (and many more)

⚠ impose some constraints on the covariance structure

What we have seen today

✔ Covariance structure

- describing a correlation structure via a graph
- correlation structure with 2 or 3 levels
- nested vs. crossed correlation factors

✔ Random effect models (2 levels)

- random intercept, random slope
- typical differences between replicates
- decomposition of the variance, e.g. within vs. between
- conditional vs. marginal predictions

✔ Comparing random effects with covariance patterns

- CS equivalent to random intercept
- random slope more flexible than CS but less than UN
- nested random intercept more flexible than CS but less than UN

Reference I

Pastes dataset

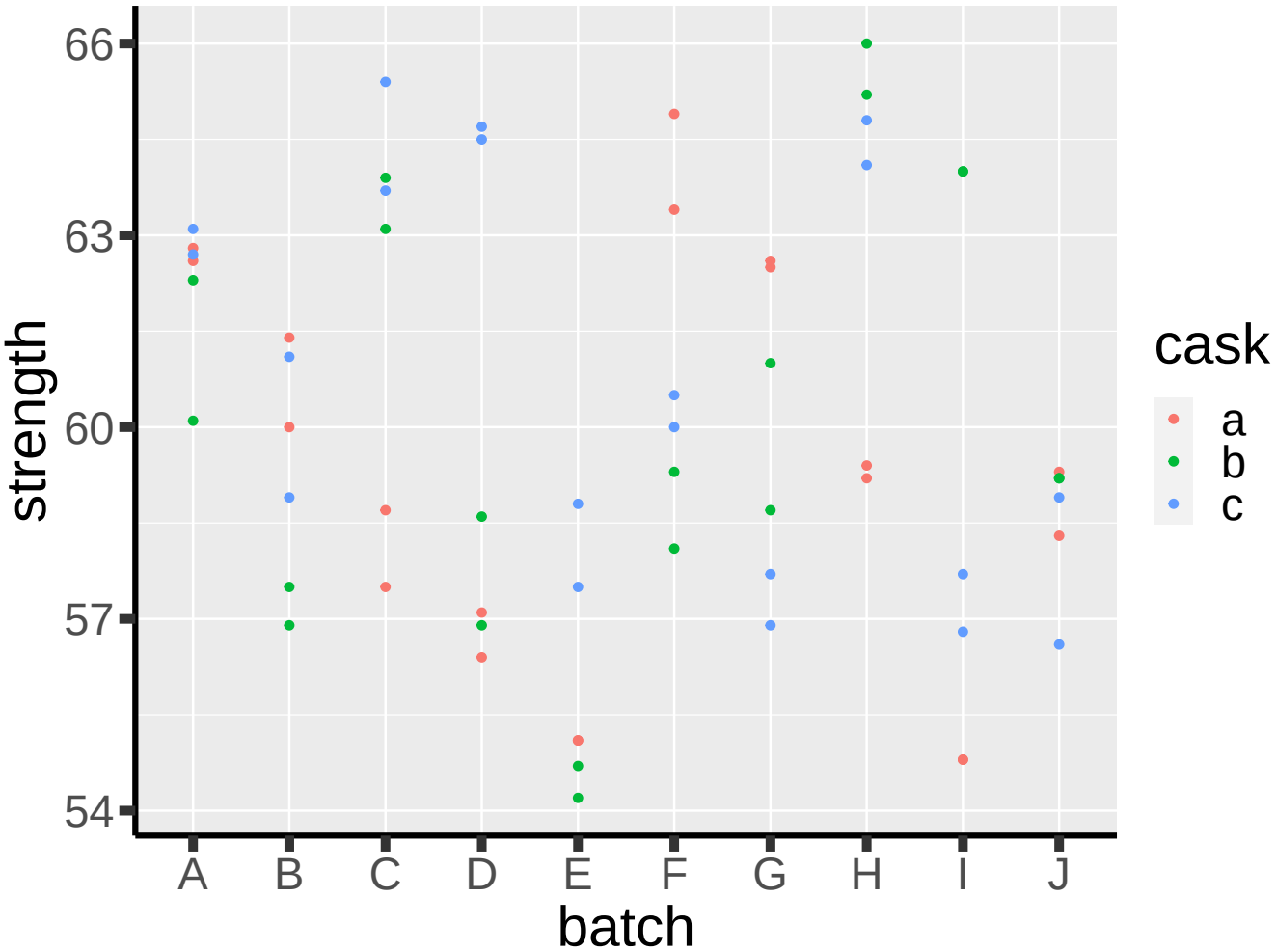
Quantify the variability in quality (`strength`) of a paste:

- several batches where analyzed (`batch`)
- in each batch several samples were taken (`sample`)

```
data("Pastes", package = "lme4")
str(Pastes)
```

```
'data.frame':        60 obs. of  4 variables:
 $ strength: num  62.8 62.6 60.1 62.3 62.7 63.1 60 61.4 57.5 56.
 $ batch   : Factor w/ 10 levels "A","B","C","D",..: 1 1 1 1 1 1
 $ cask    : Factor w/ 3 levels "a","b","c": 1 1 2 2 3 3 1 1 2 2
 $ sample  : Factor w/ 30 levels "A:a","A:b","A:c",..: 1 1 2 2 3
```

Pastes dataset (visually)



Pastes dataset (structure)

For each of the 10 batches:

- 3 samples were taken and analyzed twice

	A:a	A:b	A:c	B:a	B:b	B:c	C:a	C:b	C:c	D:a	D:b	D:c	E:a	E:b	
A	2	2	2	0	0	0	0	0	0	0	0	0	0	0	...
B	0	0	0	2	2	2	0	0	0	0	0	0	0	0	...
C	0	0	0	0	0	0	2	2	2	0	0	0	0	0	...
D	0	0	0	0	0	0	0	0	0	2	2	2	0	0	...
E	0	0	0	0	0	0	0	0	0	0	0	0	2	2	...
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
H	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...
J	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...

Nested random effects (model)

For observation $i \in \{1, 2\}$ from batch $b \in \{A, B \ldots, J\}$ and sample $s \in \{A : a, A : b \ldots, J : c\}$, we model:

$$Y_{bsi} = \alpha + u_b + v_s + \varepsilon_{bsi}$$

- α: average quality
- u_b: deviation from average quality for batch b.
  $\rightarrow$ has variance τ_b.
- v_s: deviation from average quality for sample s.
  $\rightarrow$ has variance τ_s.
- ε_{bsi}: residual error.
  $\rightarrow$ has variance σ^2.

⚠ In this model, u_b, v_s, and ε_{bsi} are independent.

Nested random effects (software)

in R :

```
e.nested <- lmer(strength ~ 1 + (1|batch/sample),
       data = Pastes)
```

in SAS:

```
proc mixed data=Pastes;
class sample batch;
model strength = / ddfm=kr vciry s;
random intercept sample
       / subject=batch;
run;
```

Nested random effects (output)

```
e.nested
```

```
Linear mixed model fit by REML ['lmerModLmerTest']
Formula: strength ~ 1 + (1 | batch/sample)
   Data: Pastes
REML criterion at convergence: 246.9907
Random effects:
 Groups       Name        Std.Dev.
 sample:batch (Intercept) 2.9041
 batch        (Intercept) 1.2874
 Residual                 0.8234
Number of obs: 60, groups:  sample:batch, 30; batch, 10
Fixed Effects:
(Intercept)
      60.05
```

Variance decomposition

Variation	Variance component	Estimate	Proportion of variance
Sample	τ_s	$2.9041^2 \approx 8.433$	78.31%
Batch	τ_b	$1.2874^2 \approx 1.657$	15.39%
Residual	σ^2	$0.8234^2 \approx 0.678$	6.29%
Total	$\tau_s + \tau_b + \sigma^2$	$.^2 + .^2 + .^2 \approx 10.769$	100%

$\rightarrow$ most of the variance comes from the sample!

Typical differences/correlation

Typical difference in strength for two observations:

- from the **same batch** and the **same sample** will be
  $[-1.96\sqrt{2 * \sigma^2}; 1.96\sqrt{2 * \sigma^2}] = [-2.28; 2.28]$

- from the **same batch** but a **different sample**
  $[-1.96\sqrt{2 * (\sigma^2 + \tau_s)}; 1.96\sqrt{2 * (\sigma^2 + \tau_s)}] = [-8.37; 8.37]$

- from a different batch (and therefore a different sample)
  $[-1.96\sqrt{2 * (\sigma^2 + \tau_s + \tau_b)}; 1.96\sqrt{2 * (\sigma^2 + \tau_s + \tau_b)}]$
  $= [-9.10; 9.10]$

Typical differences/correlation

Typical difference in strength for two observations:

- from the **same batch** and the **same sample** will be $[-1.96\sqrt{2 * \sigma^2}; 1.96\sqrt{2 * \sigma^2}] = [-2.28; 2.28]$

$\rightarrow$ correlation: $\rho_s = \frac{\tau_b + \tau_s}{\tau_b + \tau_s + \sigma^2} \approx 0.937$

- from the **same batch** but a **different sample** $[-1.96\sqrt{2 * (\sigma^2 + \tau_s)}; 1.96\sqrt{2 * (\sigma^2 + \tau_s)}] = [-8.37; 8.37]$

$\rightarrow$ correlation: $\rho_b = \frac{\tau_b}{\tau_b + \tau_s + \sigma^2} \approx 0.154$

- from a different batch (and therefore a different sample) $[-1.96\sqrt{2 * (\sigma^2 + \tau_s + \tau_b)}; 1.96\sqrt{2 * (\sigma^2 + \tau_s + \tau_b)}]$ $= [-9.10; 9.10]$

$\rightarrow$ correlation: 0

What about Σ and R?