



Faculty of Health Sciences



Analysis of mixed models with LMMstar

Stockholm biostatistics network event, June 2022

Julie Forman, Brice Ozenne

Section of Biostatistics, University of Copenhagen



Background and aims

The section of biostatistics at UCPH runs a statistical consulting service for the faculty of health science. Here we often encounter longitudinal data and other kinds of repeated measurements.

Curiously we found that available R-packages were unable to handle some of the study designs most often employed by our health science colleagues. This was our motivation for creating the LMMstar package (1st version on CRAN october 2021).

Today we will teach you about linear mixed models (LMMs) for balanced repeated measurement designs, aka mixed models for repeated measurements (MMRMs), and how to analyze them with LMMstar using health science case studies for illustration.



Outline

Part 1: Linear mixed models for repeated measurements

Practical 1

Part 2: Analysis of baseline follow-up studies

Practical 2

Part 3: Advanced topics (separate slides)

References

Supplementary: Model diagnostics



Repeated measurements

Repeated measurements refer to data where the **same outcome** is measured **several times** on the **same subjects**, e.g. on different occasions or with different treatments.

This allows us to study treatment differences and changes over time within the same subjects and the factors that influence them.

By comparing the responses of the same individual under different circumstances, we eliminate extraneous but unavoidable sources of random variation among individuals. Thus we can obtain more accurate estimates and more certain statistical conclusions.

However, since repeated measurements are typically correlated, we need dedicated models to handle them. We cannot use models which assume that observations are independent.



Examples of study designs with repeated measurements

Follow-up studies.

- ▶ Repeated measurements over time from one or more groups subjects, randomized or non-randomized.

Block designs.

- ▶ Same subjects receiving different treatments at the same time, e.g. at different locations of the body.

Crossover studies.

- ▶ Same subjects receiving different treatments in different periods of time with suitable wash-out periods inbetween.

Repeatability or reproducibility study.

- ▶ Repeated measurements with same and/or different devices to assess their mutual bias and/or precision.



Notation

We describe repeated measurements data as:

- ▶ **Subjects** $i = 1, \dots, n$.
- ▶ **Repetitions** $j = 1, \dots, k_i$.
- ▶ **Outcomes** Y_{i1}, \dots, Y_{ik_i} (from subject i).
- ▶ **Covariates** X_{ij1}, \dots, X_{ijp} (for subject i at repetition j).

In most applications the repetitions are the same for all subjects in which case we have a **balanced** study design ($k_i = k$ for all i).

Today we will only teach you models for balanced study designs.



Note about incomplete data

In practice it often happens that the data from a repeated measurements study is incomplete because some of the observations are missing.

This does not change the fact that the study design is balanced and a linear mixed models for a balanced design can still be applied.

In fact, the linear mixed model handles the missing data optimally under the assumption that it is missing at random (MAR).



The linear mixed model for repeated measurements

We assume a linear model for the outcome vectors:

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{ik} \end{pmatrix} = \begin{pmatrix} X_{i11} & X_{i12} & \dots & X_{i1p} \\ X_{i21} & X_{i22} & \dots & X_{i2p} \\ \vdots & \vdots & & \vdots \\ X_{ik1} & X_{ik2} & \dots & X_{ikp} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{ik} \end{pmatrix}$$

Note that means are modeled *similar to ordinary linear models* by including covariates called **fixed effects**. These can be continuous, categorical or of mixed type. Interactions may also be included.

Assumption: Error terms are multivariate normal $\varepsilon_i \sim \mathcal{N}(0, \Sigma_i)$.

The covariance is often assumed to be the same for all subjects, but LMMstar allows for different Σ 's in different groups.



Modeling covariances

There are multiple competing ways of modeling covariances in linear mixed models.

Random effects models are the original linear mixed models invented in the 1980s and still the most often used in practice. In these correlations are induced by subject specific random effects that are shared between the repeated measurements.

► **R-functions:** `lme4::lmer` and `nlme::lme`.

Covariance pattern models were developed in the 1990s. They allow us to model covariance more directly and flexibly than the original models. I.e. with less risk of bias due to misspecification.

► **R-packages:** `LMMstar::lmm` and `nlme::gls`.



Covariance patterns in LMMstar

Unstructured covariance, UN

$$\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \sigma_1\sigma_3\rho_{13} & \sigma_1\sigma_4\rho_{14} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 & \sigma_2\sigma_3\rho_{23} & \sigma_2\sigma_4\rho_{24} \\ \sigma_1\sigma_3\rho_{13} & \sigma_2\sigma_3\rho_{23} & \sigma_3^2 & \sigma_3\sigma_4\rho_{34} \\ \sigma_1\sigma_4\rho_{14} & \sigma_2\sigma_4\rho_{24} & \sigma_3\sigma_4\rho_{34} & \sigma_4^2 \end{pmatrix}$$

Fully flexible pattern where all variances and correlations are potentially different.

Blocked comp.sym, CS(~block)

$$\begin{pmatrix} \sigma_1^2 & \sigma_1^2\rho_1 & \sigma_1\sigma_2\rho_{12} & \sigma_1\sigma_2\rho_{12} \\ \sigma_1^2\rho_1 & \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \sigma_1\sigma_2\rho_{12} \\ \sigma_1\sigma_2\rho_{12} & \sigma_1\sigma_2\rho_{12} & \sigma_2^2 & \sigma_2^2\rho_2 \\ \sigma_1\sigma_2\rho_{12} & \sigma_1\sigma_2\rho_{12} & \sigma_2^2\rho_2 & \sigma_2^2 \end{pmatrix}$$

One variance and one correlation per block, one cross-correlation for each pair of blocks.

Compound symmetry, CS

$$\begin{pmatrix} \sigma^2 & \sigma^2\rho & \sigma^2\rho & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 & \sigma^2\rho & \sigma^2\rho \\ \sigma^2\rho & \sigma^2\rho & \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2\rho & \sigma^2\rho & \sigma^2 \end{pmatrix}$$

All variances and correlations are identical, like in a model with a random effect of subject.

Independence ID

$$\begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}$$

Covariance pattern models include ordinary linear models as a special case.



Likelihood inference

Inference is based on the likelihood function:

$$\prod_{i=1}^n \left(\frac{1}{2\pi|\Sigma(\theta)|} \right)^{\frac{k}{2}} \exp \left\{ -\frac{(y_i - X_i\beta)^T \Sigma(\theta)^{-1} (y_i - X_i\beta)}{2} \right\}$$

If $\Sigma = \Sigma_0$ is known, then the MLE for β is:

$$\hat{\beta} = \left\{ \sum_{i=1}^n (X_i^T \Sigma_0^{-1} X_i) \right\}^{-1} \sum_{i=1}^n X_i^T \Sigma_0^{-1} Y_i \quad (\text{GLS})$$

Restricted likelihood

1. Fix *any* working covariance to get an initial estimate of β .
2. Estimate the covariance from the 'likelihood' of the resulting residuals, a non-linear optimization problem.
3. Re-estimate β using this covariance.




Case: A single group follow-up study

Original Article

OBESITY BIOLOGY AND INTEGRATED PHYSIOLOGY

Obesity

Gut Mucosal Gene Expression and Metabolic Changes After Roux-en-Y Gastric Bypass Surgery

Tina Jorsal ¹, Marie M. Christensen^{1,2}, Brynjulf Mortensen¹, Eva B. Nygaard³, Chen Zhang³, Kristoffer Rigbolt³,

Study population: $n = 20$ gastric bypass patients.

Repeated measurements: 3 months before, 1 week before, 1 week after, and 3 months after surgery.

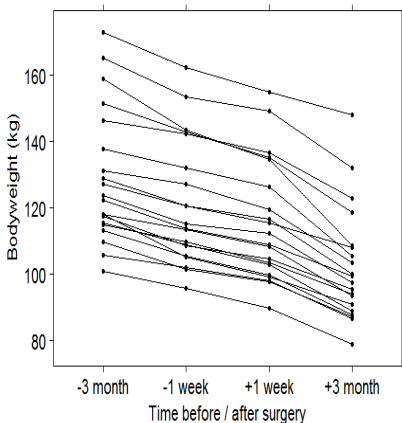
Outcomes: Bodyweight, gut hormones, and many other.

Research aims: Describe physiological changes in response to the surgery, including the pre-surgery diet.



Descriptive statistics

Spaghettiplot



Sample means and SDs

	time	obs	mean	sd
	3m before	20	129.0	20.3
	1w before	20	121.2	18.9
	1w after	20	115.7	18.3
	3m after	20	102.4	17.1

Sample correlations

	w1	w2	w3	w4
w1	1.00	0.99	0.99	0.95
w2	0.99	1.00	1.00	0.96
w3	0.99	1.00	1.00	0.97
w4	0.95	0.96	0.97	1.00



Modeling considerations

We have quantitative repeated measurements that are approx. multivariate normally distributed (possibly after transformation).

- ▶ We can use a **linear mixed model**.

We are interested in the **changes in weight over time**.

- ▶ **Fixed effect:** time (categorical, i.e. a factor in R)

How do we model the covariance?

- ▶ **We have a balanced design**, i.e. four follow-up times, which are the same for all the study participants.
- ▶ We **don't have to make any restrictive assumptions** about the covariance. We can apply an unstructured pattern.



R-syntax

The core functionality of `LMMstar` is the `lmm` function which creates the linear mixed model objects:

```
fit <- lmm(weight~time,  
           repetition=~visit|id,  
           structure="UN",  
           data=long)
```

- ▶ Use `lmm` to fit a linear mixed model with a covariance pattern.
- ▶ Outcomes and covariates are specified in a model formula, similar to ordinary linear models (`weight~time`).
- ▶ Use `repetition` to specify repeated measurements identified by time and subjects identified by `id`
- ▶ Use `structure="UN"` to choose the unstructured pattern.
- ▶ Data should be in the long format...



Wide and long data format

Data is often stored in **wide format** with one record per subject:

id	weight1	weight2	weight3	weight4	glucagonAUC1	...
2	165.2	153.4	149.2	132.0	12142.50	...
3	109.7	101.6	97.7	87.1	10321.35	...
.

However, to make linear mixed model analyses we need data to be in the **long format** with multiple records per subject:

id	visit	time	weight	glucagon.auc
2	1	3monthsBefore	165.2	12142.5
2	2	1weekBefore	153.4	14083.5
2	3	1weekAfter	149.2	10945.5
2	4	3monthsAfter	132.0	7612.5
.



Wide to long format

Data can be transformed from wide to long format in several different ways. One option is to use the reshape function:

```
long <- reshape(wide,  
  
  direction='long',  
  idvar='id',  
  varying=list(  
    c('weight1', 'weight2', 'weight3', 'weight4'),  
    c('glucagonAUC1', 'glucagonAUC2', 'glucagonAUC3', 'glucagonAUC4')  
  ),  
  v.names=c('weight', 'glucagonAUC'),  
  timevar='visit'  
  
)
```



Extractor functions for Imm objects

`summary`: Detailed model summary

`confint`: Confidence intervals and other related statistics, including standard errors and p-values.

`sigma`: Estimated covariance matrix.

`anova`: F-tests of fixed effects/linear contrasts.

`predict`: Predictions based on covariates and other outcomes.

`residuals`: Various residuals for diagnostics.

`plot`: Illustration of model fit.



Summary of fixed effects

Estimates, confidence intervals and p-values for the fixed effects (in this case changes in mean weight since baseline).

Fixed effects: `weight ~ time`

	estimate	se	df	lower	upper	p.value	
(Intercept)	128.97	4.532	18.981	119.483	138.457	<0.001	***
<code>time1weekBefore</code>	-7.73	0.697	18.974	-9.19	-6.27	<0.001	***
<code>time1weekAfter</code>	-13.27	0.839	18.969	-15.027	-11.513	<0.001	***
<code>time3monthsAfter</code>	-26.605	1.549	18.964	-29.848	-23.362	<0.001	***

Degrees of freedom were computed using a Satterthwaite approximation

Note: If you want to estimate changes in mean between other follow-up times, change the reference time with `relevel` or use `anova` to specify the relevant linear contrasts.



Summary of covariance parameters

Residual variance-covariance: unstructured

- correlation structure: ~visit

	1	2	3	4
1	1.000	0.990	0.986	0.946
2	0.990	1.000	0.997	0.959
3	0.986	0.997	1.000	0.966
4	0.946	0.959	0.966	1.000

- variance structure: ~visit

	standard.deviation	ratio
1	20.26942	1.0000000
2	18.91022	0.9329435
3	18.27535	0.9016220
4	17.05391	0.8413614

Same as summary statistics since the model is fully flexible and there are no missing observations for this outcome.



Linear mixed model vs paired t-test

Estimate (95% CI) of mean weight losses since baseline

Period	Paired t-test	lmm
	Exact degrees of freedom	Satterthwaite degrees of freedom
-3m to -1w	-7.73 (-9.19;-6.27), $P < 0.001$	-7.73 (-9.19;6.27), $P < 0.001$
-3m to +1w	-13.27 (-15.03;-11.51), $P < 0.001$	-13.27 (-15.03;-11.51), $P < 0.001$
-3m to +1m	-26.61 (-29.85;-23.36), $P < 0.001$	-26.61 (-29.85;-23.36), $P < 0.001$

We have a fully flexible model and no missing data. Hence results are identical up to a small numerical error.

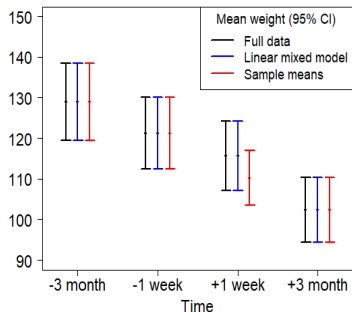
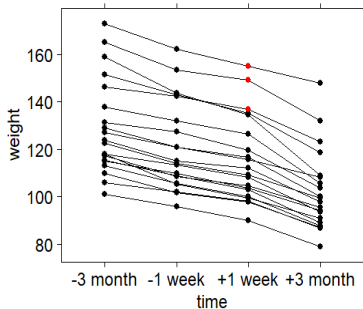
Linear mixed models handle missing data that are missing at random optimally whereas t-tests make complete case analysis.

Paired t-tests should be preferred when sample size is small ($n < 15$ per group) as the approximation of degrees of freedom in the linear mixed model may become unreliable.



Missing data

Illustrative example: Suppose three of the most obese study participants were feeling sick and didn't show up at the 3rd visit.



Sample mean at third visit is biased whereas the mixed model can predict the missing values from the neighbour observations.

Outline

Part 1: Linear mixed models for repeated measurements

Practical 1

Part 2: Analysis of baseline follow-up studies

Practical 2

Part 3: Advanced topics (separate slides)

References

Supplementary: Model diagnostics



Practical 1

Run the program `gastricbypass.R` to reproduce the analysis of the gastric bypass study.

If you have spare time, try to make a similar analysis of the second outcome in the data `glucagonAUC`.

- ▶ How do the analyses of the two outcomes compare?
- ▶ How does the covariance pattern model compare with the random effects model with a random effect of `id`?



Outline

Part 1: Linear mixed models for repeated measurements

Practical 1

Part 2: Analysis of baseline follow-up studies

Practical 2

Part 3: Advanced topics (separate slides)

References

Supplementary: Model diagnostics



Typical set-up for longitudinal data analysis

Two or more groups of subjects followed over time.

- ▶ **groups** can be randomized or non-randomized
- ▶ **time** can be calendar time, age, or duration of illness/treatment

Typical applications:

- ▶ Compare the time-evolution of a biomarker between patients and healthy controls.
- ▶ Evaluate the effect of a novel treatment in a randomized baseline follow-up study.

Do the time courses differ between the groups?



Observational vs randomized studies

In an **observational study**, we have two or more samples from inherently **different populations** we want to compare.

- ▶ Group differences may be due to confounders, so we would like to adjust for these as covariates in the analysis.
- ▶ Often we do this by including baseline covariates (and their interactions with time) in the linear mixed model.

In a **randomized study**, we have two or more samples from the **same population**, randomized to different treatments.

- ▶ Genuine group differences can only be due to treatment, so we do **not** have to worry about potential confounders (unless we have missing data).
- ▶ We gain power from making **baseline adjustment** (\rightarrow).



Case: ABETA (observational study)

Journal of Affective Disorders 277 (2020) 486–494



Contents lists available at [ScienceDirect](#)

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

The impact of a new affective episode on psychosocial functioning, quality of life and perceived stress in newly diagnosed patients with bipolar disorder: A prospective one-year case-control study

Josefine Pech^a, Morten Akhøj^b, Julie Forman^b, Lars Vedel Kessing^a, Ulla Knorr^{a,*}

Study population: $n = 87$ outpatients with bipolar disorder and $n = 44$ age and gender matched healthy controls.

Repeated measurements: at inclusion and at one-year follow-up.

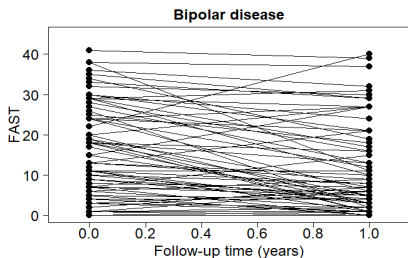
Outcomes: Functioning assessment (FAST), perceived stress (PSS), and quality of life (QoL), clinical outcomes, and other.

Research aims: Describe psycho-social functioning in patients during and after an affective episode.

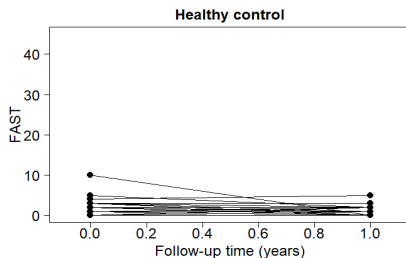


Descriptive statistics

Bipolar disorder (BD)



Healthy controls (HC)



```
> summarize(fast~year+group,
            data=long, na.rm=TRUE)
```

year	group	obs	missing	mean	sd
0	HC	44	0	1.61	1.78
1	HC	41	3	0.78	1.13
0	BD	87	0	15.58	11.50
1	BD	73	14	11.03	10.93

BD	fast0	fast1
fast0	1.000	0.308
fast1	0.308	1.000

HC	fast0	fast1
fast0	1.000	0.755
fast1	0.755	1.000



Modeling considerations

We are interested in comparing the outcomes of the groups and their evolution with time. Hence we include group, time, and the group*time interaction as fixed effects (all categorical).

Idealy, design variables (age and sex) should also be included along with potential confounders, but we omit them to save space.

We have a balanced design with two follow-up times. Hence we can apply an unstructured covariance pattern. . .

However, the patients form a more heterogeneous group and have less predictable time evolutions than the healthy controls. Hence we will need **heterogeneous covariances** with distinct parameters for each group.



Syntax

LMMstar can fit heterogeneous covariance pattern models.

```
fit2 <- lmm(fast~group*year,  
           repetition=~year|id,  
           structure=UN(~group),  
           data=long,  
           control = list(optimizer = "FS"))
```

- ▶ Use `structure=UN` with the further argument `~group` to have potentially different covariances for the two groups.
- ▶ In order to fit a heterogeneous covariance pattern, use `optimizer = "FS"` instead of the default `"gls"`.



Summary of fixed effects

Fixed effects: fast ~ group * year

	estimate	se	df	lower	upper	p.value	
(Intercept)	1.614	0.269	43.009	1.072	2.155	< 0.001	***
groupBD	13.961	1.262	93.946	11.455	16.467	< 0.001	***
year1	-0.846	0.274	43.091	-1.4	-0.293	0.00355	**
groupBD:year1	-3.643	0.951	88.699	-5.531	-1.754	< 0.001	***

Uncertainty was quantified using model-based standard errors (column se)
 Degrees of freedom were computed using a Satterthwaite approximation (c
 The columns lower and upper indicate a 95% confidence interval for each

Substantially worse functioning in the bipolar patients at entry.
 Improvement in both groups over time, but more in the patient
 group than in the healthy control group.



Summary of covariance parameters

Residual variance-covariance: stratified unstructured

- correlation structure: ~year:group - 1

\$HC			\$BD		
	0	1		0	1
0	1.000	0.303	0	1.000	0.753
1	0.303	1.000	1	0.753	1.000

- variance structure: ~0 + group + year:group

	standard.deviation	ratio
sigma.0:HC	1.781245	1.0000000
sigma.0:BD	11.500892	1.0000000
sigma.1:HC	1.127808	0.6331572
sigma.1:BD	10.889847	0.9468697

Estimates are similar to the summary statistics, but not exactly the same since there are missing observations for this outcome.



Homogeneous vs heterogeneous covariances

Unrecognized heterogeneity doesn't lead to biased estimates unless there are severe issues with missing data. Standard errors, however, will be biased unless heterogeneity is accounted for.

Effect	Heterogeneous UN-Pattern		Homogeneous UN-Pattern		Bias
	Estimate	SE	Estimate	SE	
Baseline mean BD	15.6	1.23	15.6	1.01	-18%
Baseline mean HC	1.6	0.27	1.6	1.42	+425%
Difference	14.0	1.26	14.0	1.75	+39%
Change at follow-up BD	-4.5	0.91	-4.5	0.74	-19%
Change at follow-up HC	-0.9	0.27	-0.9	1.00	+270%
Difference	-3.6	0.95	-3.6	1.24	+31%

SE too small SE too large.



Case study: CKD (randomized study)

OPEN ACCESS Freely available online



Eplerenone Attenuates Pulse Wave Reflection in Chronic Kidney Disease Stage 3–4 - A Randomized Controlled Study

Lene Boesby^{1*}, Thomas Elung-Jensen², Svend Strandgaard¹, Anne-Lise Kamper²

¹Department of Nephrology, Herlev Hospital, University of Copenhagen, Herlev, Denmark, ²Department of Nephrology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

Study population: $n = 51$ patients newly diagnosed with chronic kidney disease randomized to Eplerenone or standard care.

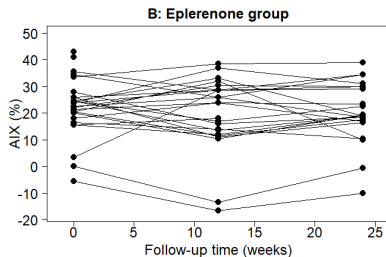
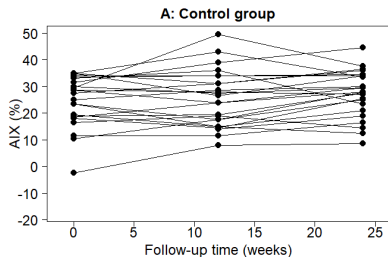
Repeated measurements: at 0 weeks (baseline), 12 weeks (safety) and 24 weeks (end point) after treatment initiation.

Outcomes: Augmentation index (aix), pulse wave velocity, other.

Research aims: Evaluate effect of Eplerenone vs standard care.



Descriptive statistics



time	group	obs	missing	mean	sd
0	A	24	1	24.6	9.38
12	A	24	1	25.3	10.60
24	A	24	1	27.3	8.70
0	B	26	0	22.3	11.11
12	B	24	2	19.9	13.70
24	B	22	4	20.4	11.43

A	aix0	aix12	aix24
aix0	1.00	0.79	0.76
aix12	0.79	1.00	0.80
aix24	0.76	0.80	1.00

B	aix0	aix12	aix24
aix0	1.00	0.68	0.73
aix12	0.68	1.00	0.82
aix24	0.73	0.82	1.00



Modeling considerations

Data looks as from an observational study with two groups followed over time, but in this case groups are randomized.

The baseline measurement is taken before treatment initiation. Since we have two random samples from the exact same population, their mean parameters must be identical.

We can estimate the treatment effect as the difference in means at follow-up. Since baseline means are identical, this is the same as the difference in *change in mean from baseline to follow-up*.

Similarly, the variances for the two groups should be identical at baseline and potentially different at follow-up.



Syntax

1. Define a time-varying treatment variable:

```
long$treat <- long$allocation
long$treat[long$time==0] <- "A"

table(long$time, long$treat)
```

	A	B
0	51	0
12	25	26
24	25	26

2. ... and its interaction with time:

```
long$time.treat <- interaction(long$time, long$treat)
```

3. Leave out the main effect of treat in the model formula:

```
fitcc <- lmm(aix~time+time:treat,
            repetition=~time.treat|id,
            structure="UN",
            data=long,
            control = list(optimizer = "FS"))
```



Summary of fixed effects

Fixed effects: aix ~ time + time:treat

	estimate	se	df	lower	upper	p.value	
(Intercept)	23.255	1.443	49.909	20.356	26.154	<0.001	***
time12	1.299	1.364	23.051	-1.521	4.12	0.3506	
time24	3.476	1.248	25.836	0.909	6.042	0.0099	**
time12:treatB	-2.158	2.52	36.336	-7.267	2.951	0.3974	
time24:treatB	-4.069	2.047	38.905	-8.211	0.072	0.0539	.

Uncertainty was quantified using model-based standard errors (column se)
 Degrees of freedom were computed using a Satterthwaite approximation (c
 The columns lower and upper indicate a 95% confidence interval for each

We find a borderline significant improvement i AIX with
 Eplerenone compared to control at final follow-up (P=0.054).



Summary of covariance parameters

```
- correlation structure: ~time.treat
      0.A  12.A  24.A  12.B  24.B
0.A  1.000 0.810 0.786 0.694 0.735
12.A 0.810 1.000 0.811    NA    NA
24.A 0.786 0.811 1.000    NA    NA
12.B 0.694    NA    NA 1.000 0.817
24.B 0.735    NA    NA 0.817 1.000
```

```
- variance structure: ~time.treat
      standard.deviation    ratio
0.A          10.276029 1.0000000
12.A          11.056805 1.0759804
24.A           9.054361 0.8811148
12.B          13.957917 1.3582988
24.B          11.391086 1.1085105
```

Some of the correlations are unidentifiable, since no one gets both treatment A and B. This is harmless.



Other approaches to randomized follow-up studies

Vickers & Altman: *Analysing controlled clinical trials with baseline and follow up measurements*, BMJ **323**, 1123–1124, (2001).

Considers three classical approaches:

1. Two-sample t-test based on outcome at follow-up.
2. Two-sample t-test based on change from baseline to follow-up
3. ANCOVA including baseline as a covariate.

$$Y_{\text{follow-up}} = \beta_0 + \beta_1 \cdot I\{\text{treat}\} + \beta_2 \cdot Y_{\text{baseline}} + \varepsilon$$

Conclusion: ANCOVA is always most efficient.



cLMM vs ANCOVA

The constrained linear mixed model (cLMM):

- ▶ aka constrained repeated measures or cMMRM.
- ▶ Include baseline as outcome in a constrained linear mixed model where **group means are identical at baseline**.

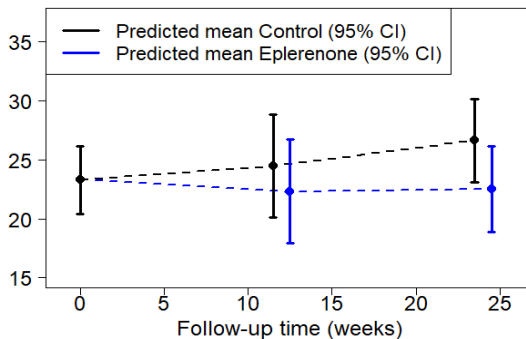
Merits:

- ▶ Equivalent to ANCOVA with no missing data.
- ▶ By default handles missing data (MAR) optimally, ANCOVA by default makes complete case analysis.
- ▶ Handles heterogeneous covariances more easily.
- ▶ Handles multiple follow-up measurements more easily.
- ▶ Model fit is easier to visualize.



Predicted values

Estimated AIX (%) for the average patient in the population had he/she received either Eplerenone or standard care.



Note: In the ANCOVA the predicted values are subject specific because they depend on the value of the outcome at baseline.



Outline

Part 1: Linear mixed models for repeated measurements

Practical 1

Part 2: Analysis of baseline follow-up studies

Practical 2

Part 3: Advanced topics (separate slides)

References

Supplementary: Model diagnostics



Practical 2

Run the programs `abeta.R` and `ckd.R` to reproduce the analyses of the two case studies.

If you have spare time, try to compare the results you get from:

- ▶ A constrained vs unconstrained model for the mean.
- ▶ A homogeneous vs heterogeneous covariance pattern.



Outline

Part 1: Linear mixed models for repeated measurements

Practical 1

Part 2: Analysis of baseline follow-up studies

Practical 2

Part 3: Advanced topics (separate slides)

References

Supplementary: Model diagnostics



Outline

Part 1: Linear mixed models for repeated measurements

Practical 1

Part 2: Analysis of baseline follow-up studies

Practical 2

Part 3: Advanced topics (separate slides)

References

Supplementary: Model diagnostics



References

Classical references:

- ▶ G.M. Fitzmaurice, N.M. Laird & J.H. Ware :
Applied Longitudinal Analysis (2nd edition),
John Wiley & sons, 2011
- ▶ J. Pinheiro & D. Bates :
Mixed-effects models in S and S-PLUS.
Springer science & business media, 2006.

Lecture notes and tutorials from our ph.d. course for health science researchers are available at:

- ▶ <https://absalon.ku.dk/courses/47665>



References

cLMM vs ANCOVA:

Liu et al: *Should baseline be a covariate or dependent variable in analyses of change from baseline in clinical trials?*,
Statistics in Medicine **28**, 2509–2530, (2009):

Coffman et al: *To condition or not condition? Analysing change in longitudinal randomised controlled trials*,
BMJ open **6.12**, e013096, (2016).

Wan: *Analyzing pre-post randomized studies with one post-randomization score using repeated measures and ANCOVA models*,
Statistical methods in medical research **28**, 2952-2974, (2019).



Outline

Part 1: Linear mixed models for repeated measurements

Practical 1

Part 2: Analysis of baseline follow-up studies

Practical 2

Part 3: Advanced topics (separate slides)

References

Supplementary: Model diagnostics



Diagnostics for linear mixed models

$$Y_i = X_i\beta + \varepsilon_i, \quad \text{where } \varepsilon_i \sim \mathcal{N}(0, \Sigma_i)$$

We can estimate the error terms by the residuals

$$r_i = Y_i - X_i\hat{\beta} \approx \varepsilon_i$$

and use them for model diagnostics like in a linear model:

1. Plots of residuals against predicted values/covariates.
2. QQplots of residuals to assess normal distribution.

However: Contrary to ordinary linear models, the residuals will be correlated and may have heterogeneous variances.

- ▶ Hence we ought to standardize them first. . .



Standardized residuals

Pearson residuals

$$r_i^* = \text{diag}(\hat{\Sigma}_i)^{-1/2}(Y_i - X_i\hat{\beta})$$

Divide by estimated standard deviations, variance ≈ 1 .

Studentized residuals

$$\tilde{r}_i^* = \text{diag}(\hat{\Sigma}_i^{(-i)})^{-1/2}(Y_i - X_i\hat{\beta}^{(-i)})$$

Jackknife estimates leaving out the i 'th subject makes it easier to detect outliers. Very similar to the Pearson residuals if sample size is large and there aren't many outliers.

However: The residuals will still be correlated. . .



De-correlated residuals

Scaled residuals

$$r_i^{**} = L_i^{-1/2}(Y_i - X_i\hat{\beta}) \approx \mathcal{N}(0, I)$$

where L_i is a lower triangular matrix such that $L_i L_i^T = \hat{\Sigma}_i$.

It can be shown that the scaled residuals are the pearson residuals from the ordinary linear models dynamically predicting each consecutive outcome based on the previous outcomes and the covariates in the linear mixed model:

$$\begin{aligned} r_{i1}^{**} &= r_{i1}^* \\ r_{ij+1}^{**} &= \frac{Y_{ij+1} - \hat{E}[Y_{ij+1}|Y_{i1}, \dots, Y_{ij}]}{\sqrt{\hat{V}[Y_{ij+1}|Y_{i1}, \dots, Y_{ij}]}} \end{aligned}$$



Diagnostics in LMMstar

Use the residuals function to obtain:

`fitted`: Predicted values based on covariates.

`r.response`: Ordinary residual.

`r.pearson`: Pearson residual.

`r.studentized`: Studentized residual.

`r.scaled`: De-correlated residual.

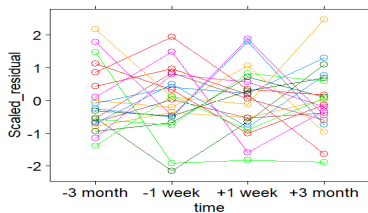
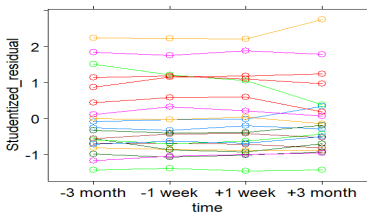
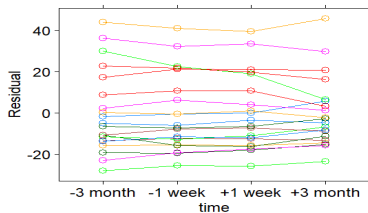
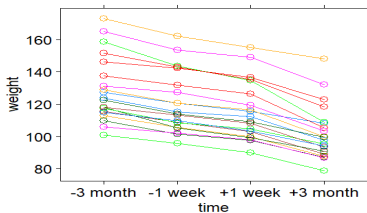
```
diag <- residuals(fit, type='all', keep.data=TRUE)
```

id	visit	weight	fitted	r.response	r.studentized	r.scaled
1	1	127.2	128.970	-1.770	-0.08959219	-0.08732367
1	2	120.7	121.240	-0.540	-0.02929782	0.40462002
1	3	115.5	115.700	-0.200	-0.01122800	0.21188023
1	4	108.1	102.365	5.735	0.34502276	1.29889231
2	1	165.2	128.970	36.230	1.83385604	1.78742177



Residuals

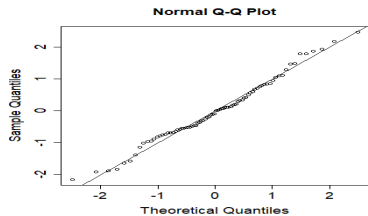
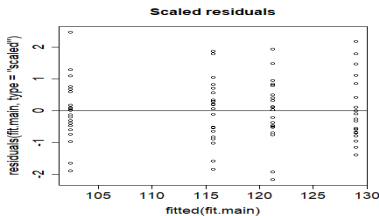
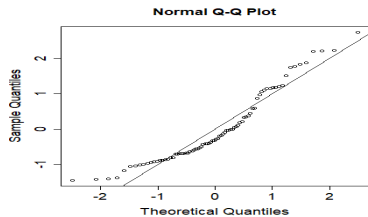
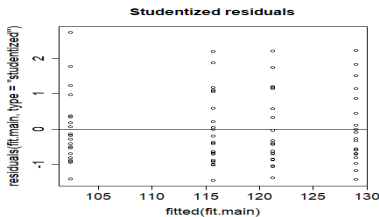
Original data, response, studentized, and scaled residuals.



Scaled residuals are independent if the model is well specified.

Diagnostics

Residual- and QQplots of studentized and scaled residuals.



Studentized residuals indicate some skewness in the marginal distribution.

How important is the normal distribution?

The **standard model assumption** is that outcomes follow a **multivariate normal distribution**.

BUT: Linear mixed models are **robust** (the central limit theorem).

- ▶ As long as the linear model for the mean is correct and the covariance is well specified.
- ▶ If sample size is not too small.
- ▶ If there aren't too many missing observations.
- ▶ If the distribution of the data is not too skew.

Highly skew data should be transformed.

