Last time: Multi- and Univariate repeated measurements ANOVA

So far, when dealing with the statistical analysis of repeated measurements, we have considered:

Multivariate repeated measurements ANOVA:

- No missing observations and less than 6-7 time points: **Exact tests**
- Missing observations and less than 6-7 time points: **Approximate tests**

Univariate repeated measurements ANOVA:

- No missing observations and more than 6-7 time points: **Exact tests**
- Missing observations and more than 6-7 time points: **Approximate tests**

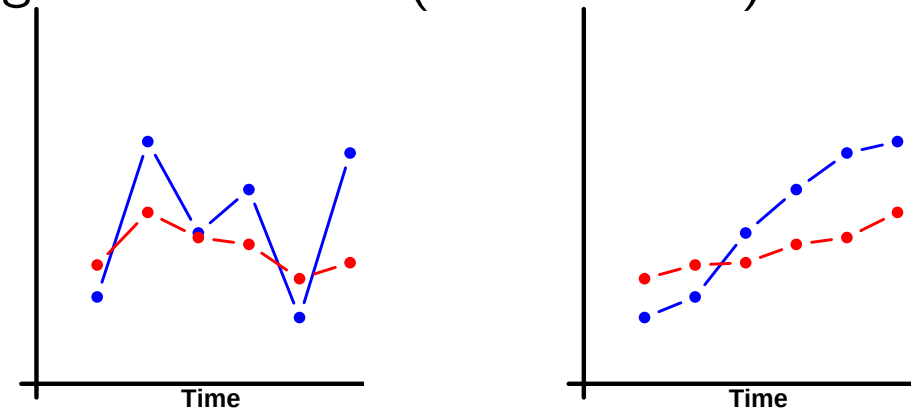
In none of the above mentioned methods is the **ordering of the time-points** taken into account.

Today: Time as a covariate and more complex designs

- Example: Growth of preadolescent girls
- The Random Coefficient Model (RCM)
- The Random Coefficient Model in Stata
- Example: Orthodontic measurement over time for boys and girls
- Example: Diet and plasma glucose for diabetic patients
- Analysis of summary statistics
- More complex designs: Cross-over trial with repeated measurements

Time as a covariate: Taking the ordering into account

Consider the following two data sets (mean curves):



- In fact the observations are **identical**. The only difference is that the time-points have been interchanged.
- An analysis based on the multivariate or univariate repeated measurements ANOVA would result in **identical tests** regarding the group difference (here there is no difference between the groups).
- For the second data set, however, if we take the ordering into account it is not hard to imagine that there is a difference between the two groups. To take the ordering of the time-points into account we need to **describe the relationship** between the response and time.

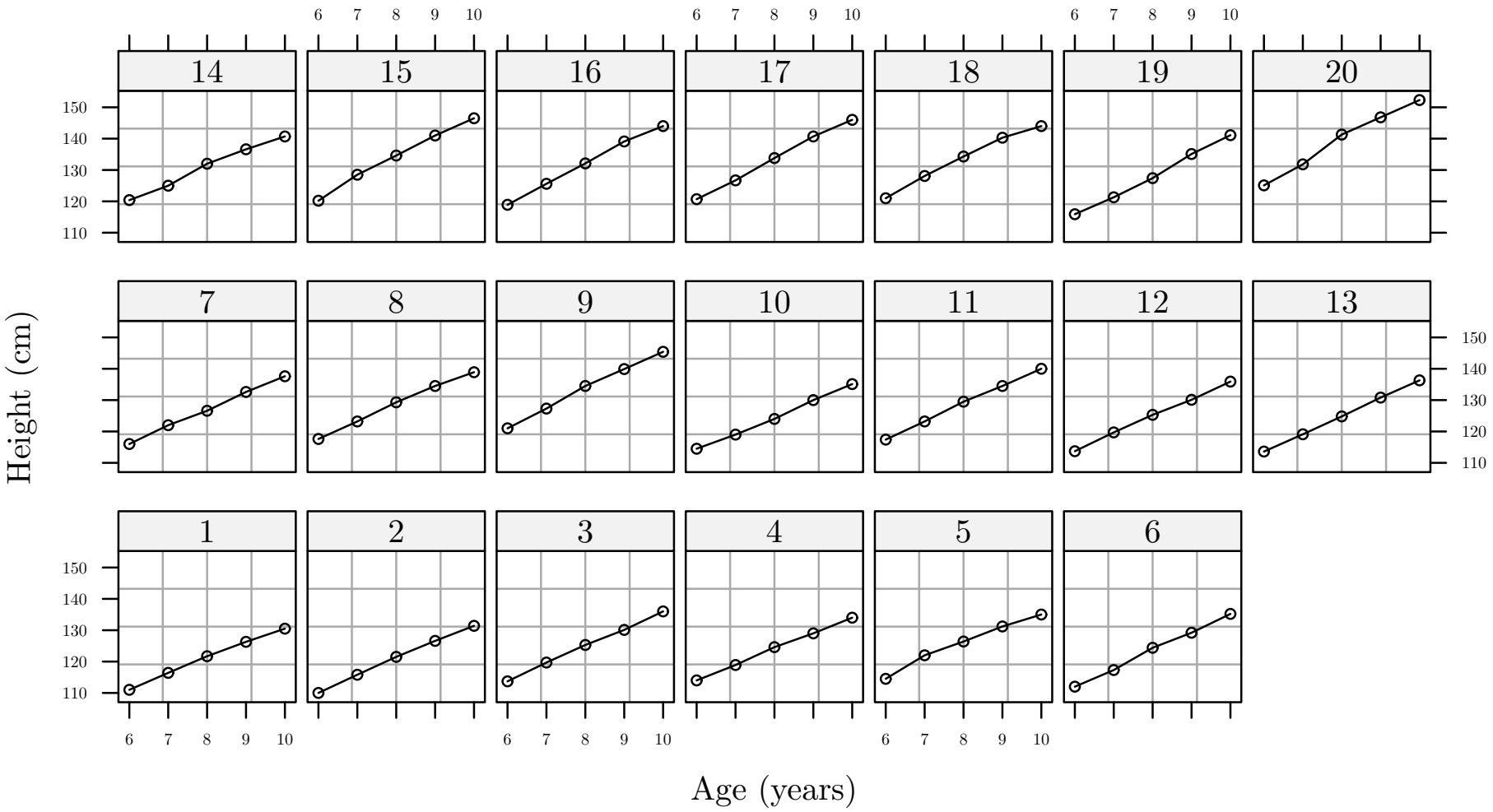
Example: Growth of preadolescent girls

We consider growth of preadolescent girls and how it depends on their mothers height. For 20 girls we have measurements of the height (cm) at 6, 7, 8, 9, and 10 years of age along with their mothers height classified as **short** (< 155 cm), **medium** ($155 - 165$ cm), or **tall** (> 165 cm).

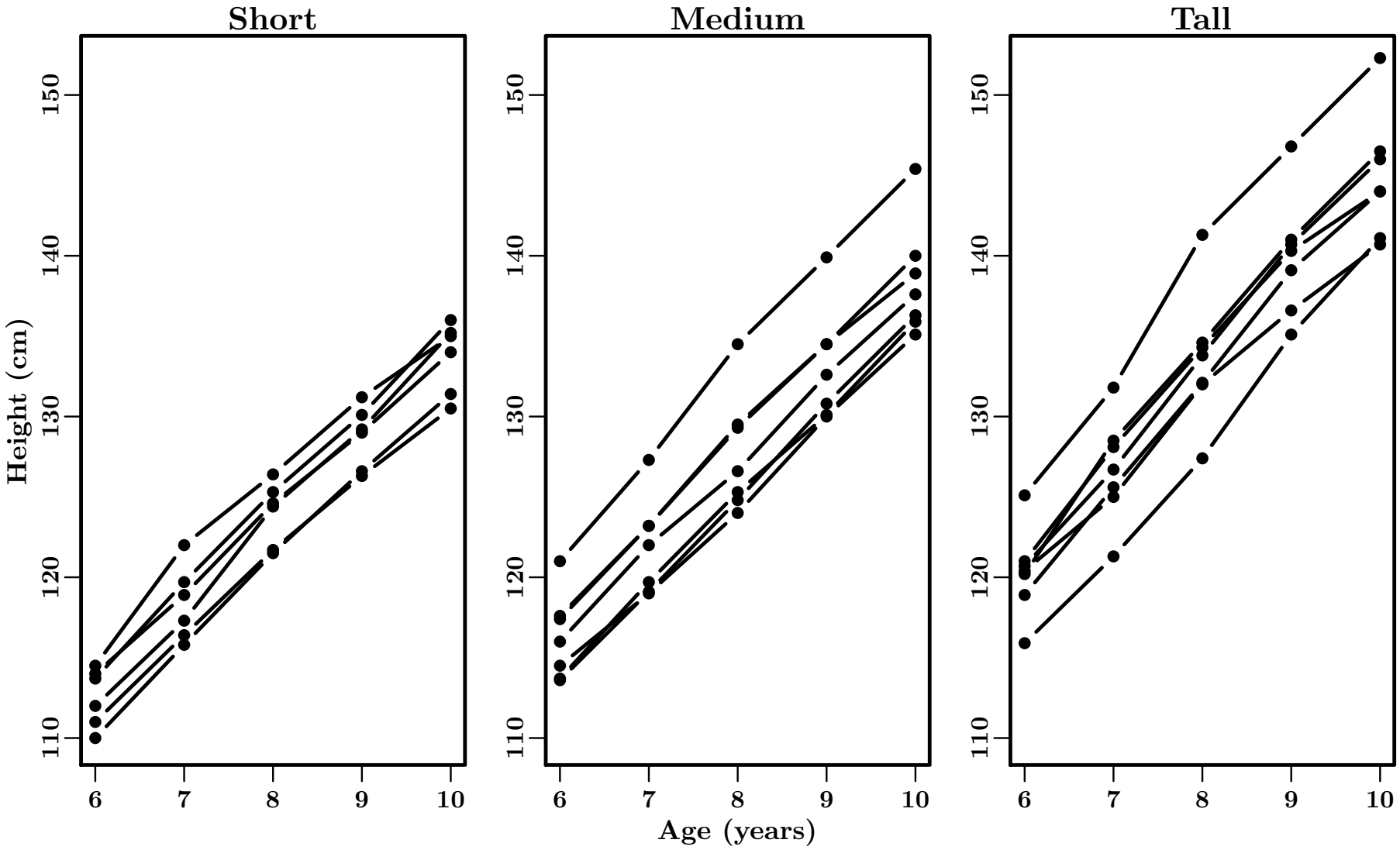
Mother	Girl	6 years	7 years	8 years	9 years	10 years
Short	1	111.0	116.4	121.7	126.3	130.5
	⋮	⋮	⋮	⋮		
Medium	7	116.0	122.0	126.6	132.6	137.6
	⋮	⋮	⋮	⋮		
Tall	14	120.4	125.0	132.0	136.6	140.7
	⋮	⋮	⋮	⋮		

Question: Does the growth of preadolescent girls depend on their mothers height?

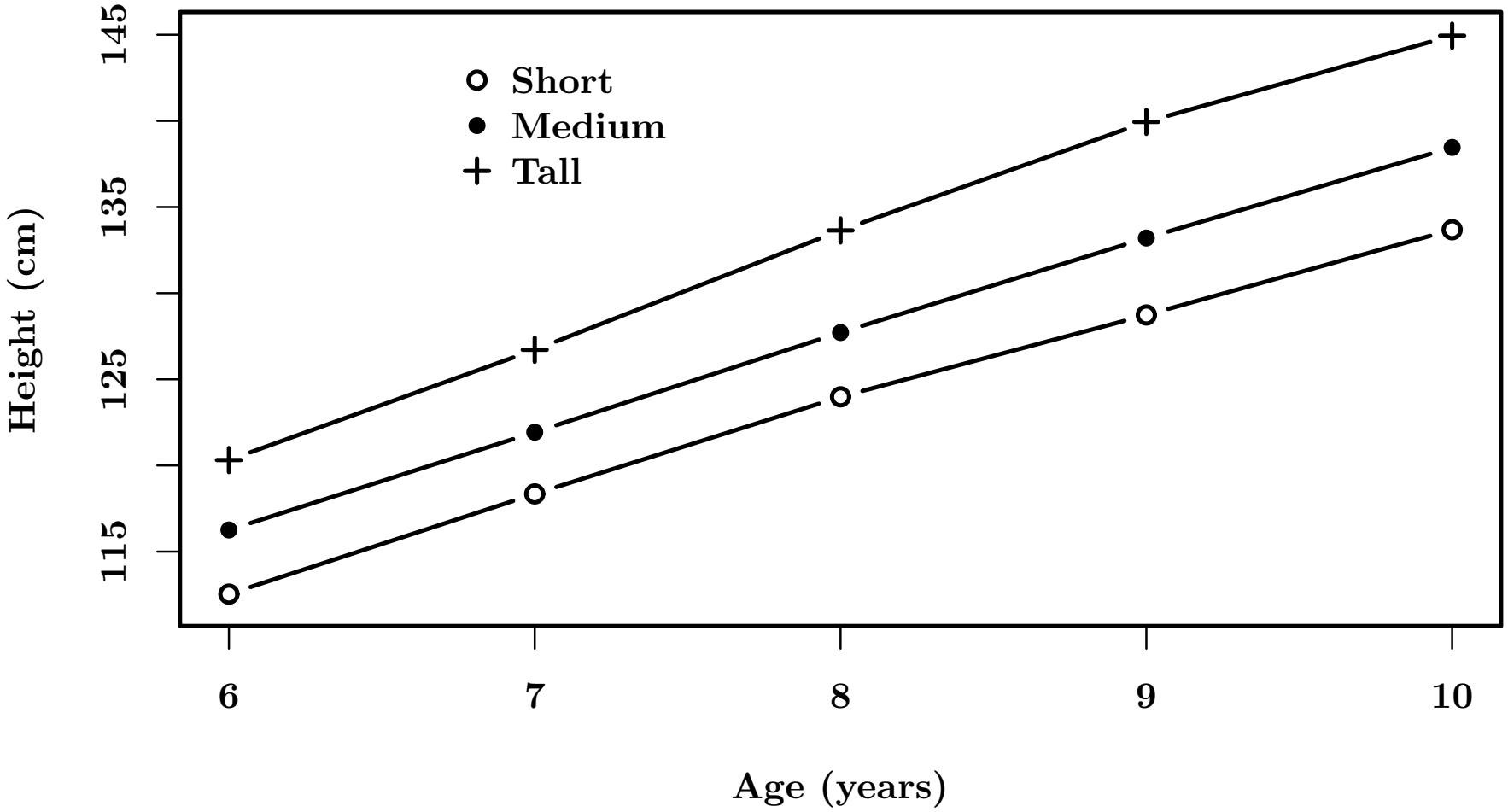
Growth of preadolescent girls: Individual curves



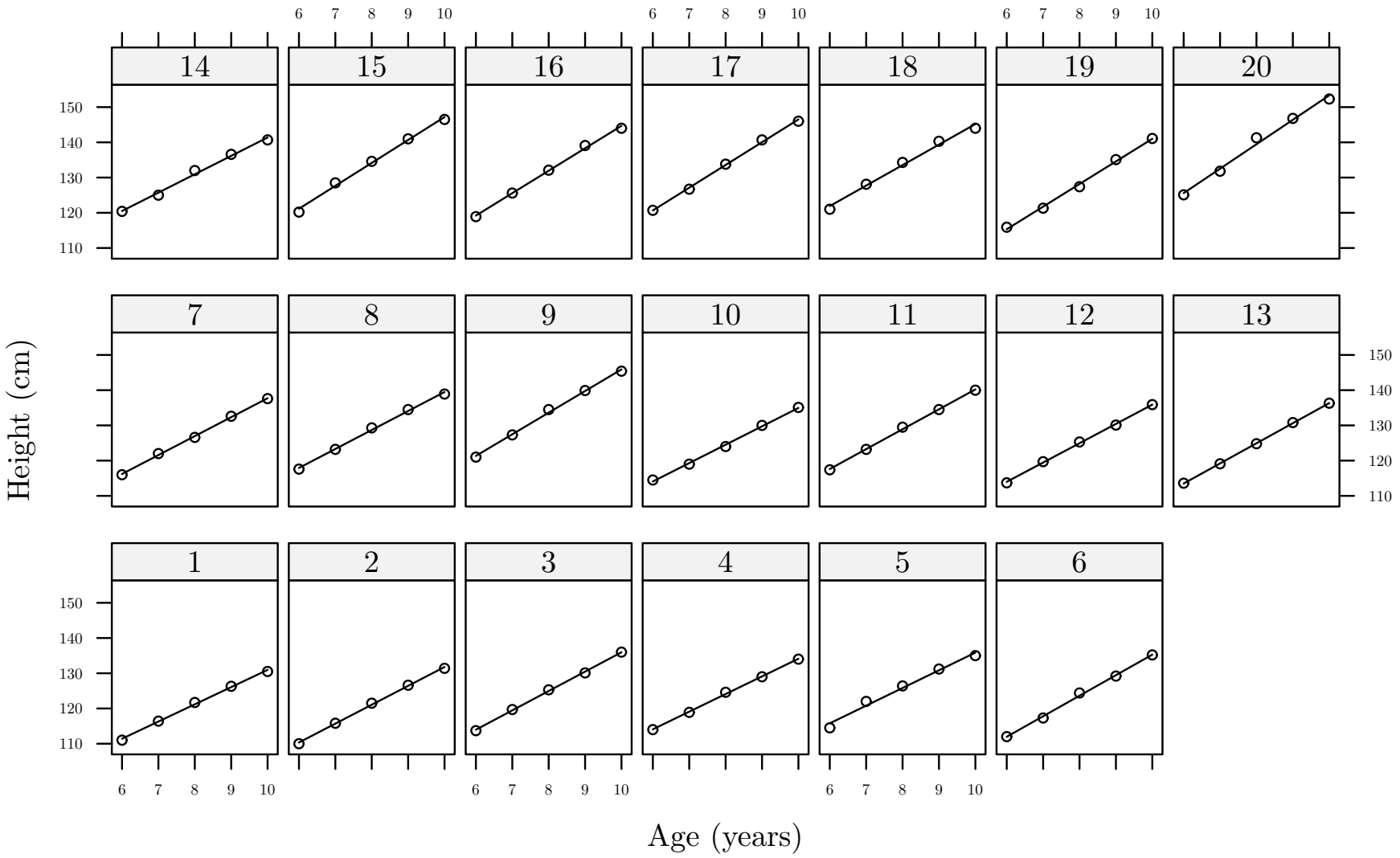
Growth of preadolescent girls: Grouped individual curves



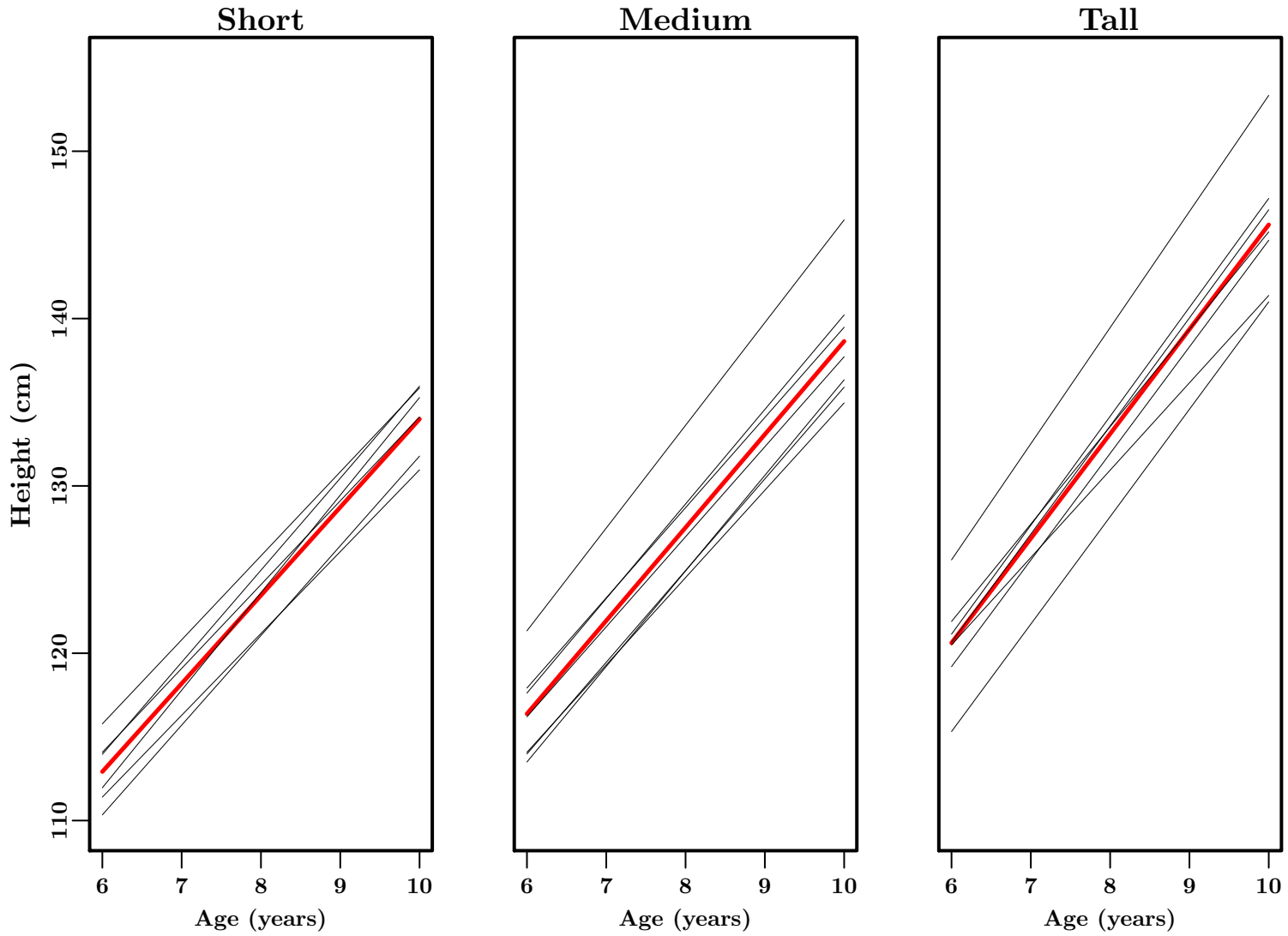
Growth of preadolescent girls: Mean curves



Growth of girls: Linear regression for each subject



Illustrating the random coefficient model



Assumptions

A **linear regression model** with randomly varying slopes and intercepts:

$$\texttt{height} = \alpha + \beta \cdot \texttt{age} + e$$

Random intercept:

$\alpha =$ normal with mean $\alpha_{\texttt{S}}$ (for girls with short mothers), and sd σ_α

Random slope:

$\beta =$ normal with mean $\beta_{\texttt{S}}$ (for girls with short mothers), and sd σ_β

Population parameters:

- **Short:** $\alpha_{\texttt{S}}$, $\beta_{\texttt{S}}$
- **Medium:** $\alpha_{\texttt{M}}$, $\beta_{\texttt{M}}$
- **Tall:** $\alpha_{\texttt{T}}$, $\beta_{\texttt{T}}$

Variation around the line:

e normal with mean 0 and standard deviation σ

Comments

We could just estimate the intercept and slope for each subject and compare the groups based on these estimates:

- That would constitute an analysis based on **summary measures** (something we will return to later today).
- We would probably want to compare the groups using a MANOVA model since the estimated intercept and slope are highly correlated (or maybe an ANOVA model based only on the slopes).
- The estimates for the population parameters based on the random coefficient model will be identical to the means of the individual intercepts and slopes, but **the standard errors will be (potentially much) bigger**.
- The random coefficient model uses **all the data** directly to compare the groups, whereas the analysis based on summary statistics excludes a lot of information by forgetting that the intercepts and slopes are estimated, with all the uncertainty that implies, when comparing the groups.

Growth of preadolescent girl: Parameter estimates

The estimated slopes and intercepts for the three populations:

$$\begin{aligned}
\hat{\alpha}_{\mathtt{S}} &= 81.3 \quad [78.5, 84.1], & \hat{\beta}_{\mathtt{S}} &= 5.3 \quad [4.9, 5.6] \\
\hat{\alpha}_{\mathtt{M}} &= 83.0 \quad [80.4, 85.6], & \hat{\beta}_{\mathtt{M}} &= 5.6 \quad [5.2, 5.9] \\
\hat{\alpha}_{\mathtt{T}} &= 83.1 \quad [80.5, 85.7], & \hat{\beta}_{\mathtt{T}} &= 6.2 \quad [5.9, 6.6]
\end{aligned}$$

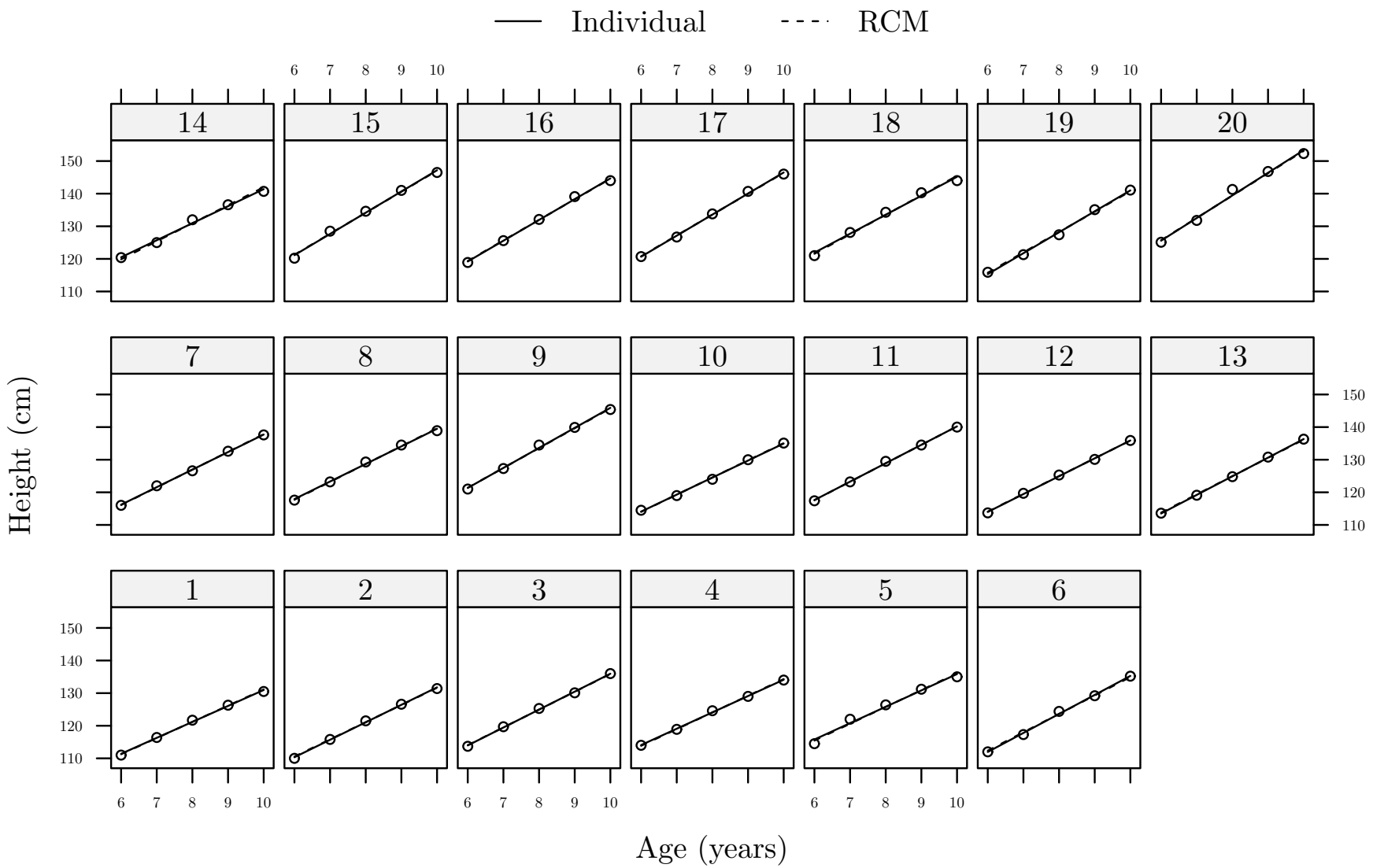
A **test for equal slopes** gives a p-value of 0.0001 and so we conclude that there is clear evidence in the data against the hypothesis of equal **growth** in the three populations.

The hypothesis of **equal growth** for girls with short and medium mothers is accepted with a p-value of 0.21, whereas the growth is significantly bigger for girls with tall mothers ($p = 0.003$ compared to medium mothers).

The standard deviation estimates are:

$$\hat{\sigma}_{\alpha} = 2.76, \qquad \hat{\sigma}_{\beta} = 0.37, \qquad \hat{\sigma} = 0.69$$

Growth of preadolescent girl: Model fit



Model validation: Observed and expected standard deviations and correlations

Observed:

Short

$$\begin{pmatrix} 1.81 & & & & \\ 0.93 & 2.32 & & & \\ 0.94 & 0.93 & 1.98 & & \\ 0.90 & 0.92 & 0.99 & 1.93 & \\ 0.77 & 0.72 & 0.91 & 0.92 & 2.23 \end{pmatrix}$$

Medium

$$\begin{pmatrix} 2.66 & & & & \\ 0.99 & 2.99 & & & \\ 0.97 & 0.99 & 3.68 & & \\ 0.98 & 0.99 & 0.99 & 3.52 & \\ 0.96 & 0.97 & 0.99 & 0.99 & 3.51 \end{pmatrix}$$

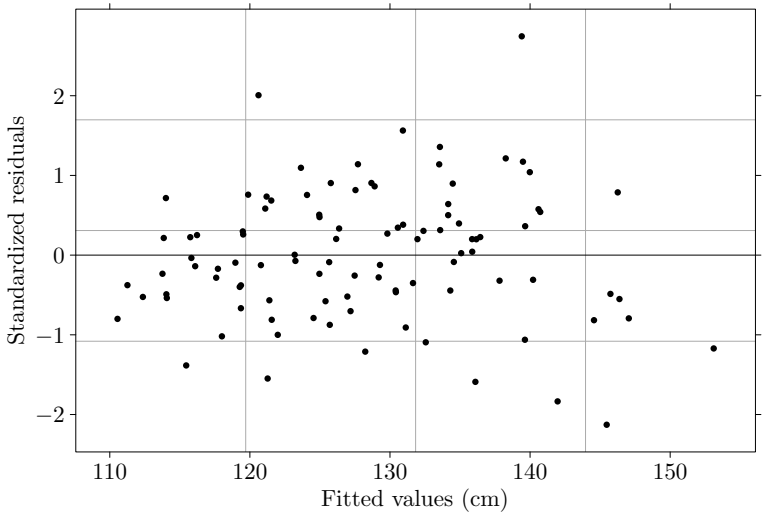
Tall

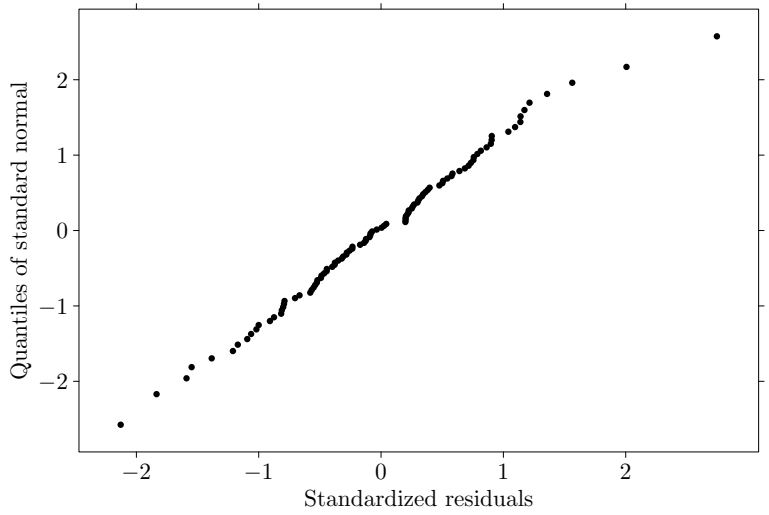
$$\begin{pmatrix} 2.74 & & & & \\ 0.94 & 3.28 & & & \\ 0.97 & 0.98 & 4.16 & & \\ 0.90 & 0.95 & 0.97 & 3.75 & \\ 0.82 & 0.88 & 0.92 & 0.98 & 3.92 \end{pmatrix}$$

Expected:

$$\begin{pmatrix} 2.75 & & & & \\ 0.93 & 2.90 & & & \\ 0.92 & 0.94 & 3.08 & & \\ 0.90 & 0.93 & 0.95 & 3.30 & \\ 0.87 & 0.92 & 0.94 & 0.96 & 3.54 \end{pmatrix}$$

Growth of preadolescent girl: Residuals and conclusions





Conclusion: No clear deviations from the random coefficient model.

Our best estimate is that **girls with tall mothers** on average grow

$$6.2 \text{ cm/year}, \qquad 95\% - \text{CI}: \ 5.9 - 6.6 \text{ cm/year}$$

when between the ages of 6 and 10 years.

That is significantly more than girls with short or medium-sized mothers.

Growth of girl: Conclusions regarding the height

We saw that the intercept estimates could not be interpreted. We get the same model for the data if we subtract 6 from all the ages and consider:

$$\texttt{height} = \alpha + \beta \cdot (\texttt{age} - 6) + e$$

This does not change the estimates for the slopes, but the intercept estimates become:

$$\begin{aligned} \hat{\alpha}_{\texttt{S}} &= 112.9 \quad [110.6, 115.3] \\ \hat{\alpha}_{\texttt{M}} &= 116.6 \quad [114.2, 118.5] \\ \hat{\alpha}_{\texttt{T}} &= 120.6 \quad [118.5, 122.8] \end{aligned}$$

We can interpret the estimates as the expected height of girls in the different groups at 6 years of age.

Conclusion: The growth is different for girls with tall mothers, but already at the age of 6 years the girls with tall mother are significantly taller than girls with medium-sized mothers ($p = 0.003$) who are significantly taller than girls with short mothers ($p = 0.022$).

Growth of preadolescent girl: The analysis in Stata

In Stata you can use `mixed` to perform the analyzes:

```
mixed height bn.mother bn.mother#c.age, nocons || ///
  girl: age, cov(un) reml
```

In Stata you can test the hypothesis of equal slopes in all three groups in the following way:

```
mixed height bn.mother bn.mother#c.age, nocons || ///
  girl: age, cov(un) mle
estimates store model1

mixed height bn.mother c.age, nocons || girl: age, cov(un) mle
estimates store model2

lrtest model1 model2
```

Growth of preadolescent girl: Output from Stata

height		Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
mother	Short	81.31	1.339289	60.71	0.000	78.68504	83.93496
	Medium	82.97428	1.239941	66.92	0.000	80.54404	85.40452
	Tall	83.12286	1.239941	67.04	0.000	80.69262	85.55309
mother#c.age	Short	5.268333	.1736548	30.34	0.000	4.927976	5.60869
	Medium	5.567143	.1607731	34.63	0.000	5.252034	5.882253
	Tall	6.248572	.1607731	38.87	0.000	5.933462	6.563681

Random-effects Parameters		Estimate	Std. Err.	[95% Conf.	Interval]
girl: Unstructured	sd(age)	.3651428	.0858106	.2303698	.5787619
	sd(_cons)	2.760457	.6766459	1.707394	4.463015
	corr(age,_cons)	-.4424586	.2556303	-.7998775	.1466397
	sd(Residual)	.6899767	.062986	.5769401	.8251599

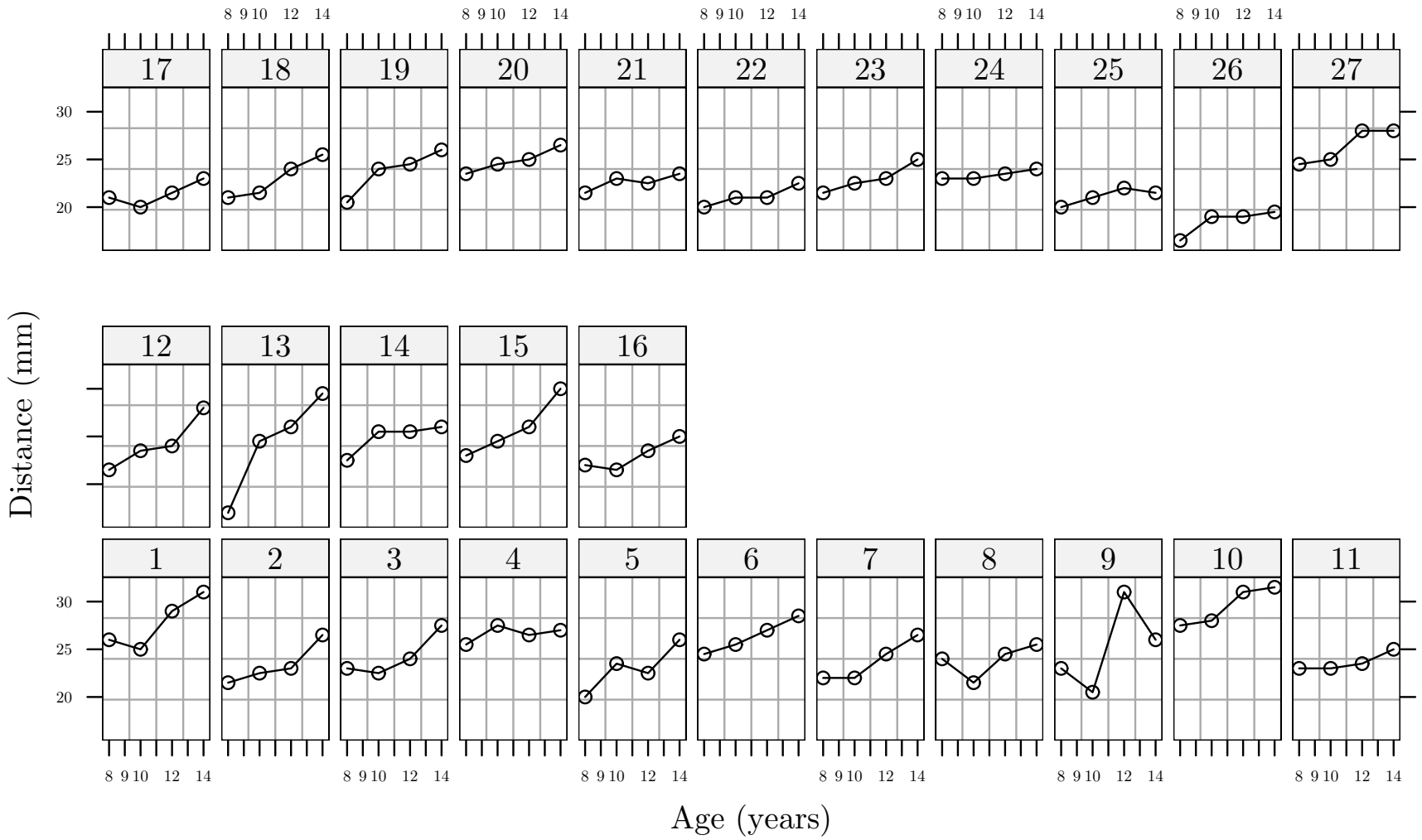
Example: Orthodontic measurements for boys and girls

We consider data on two groups of children (16 boys and 11 girls). At ages 8, 10, 12, and 14 years, the distance (mm) from the center of the pituitary gland to the pterygomaxillary fissure was measured. Data are from **Potthoff & Roy (1964)**.

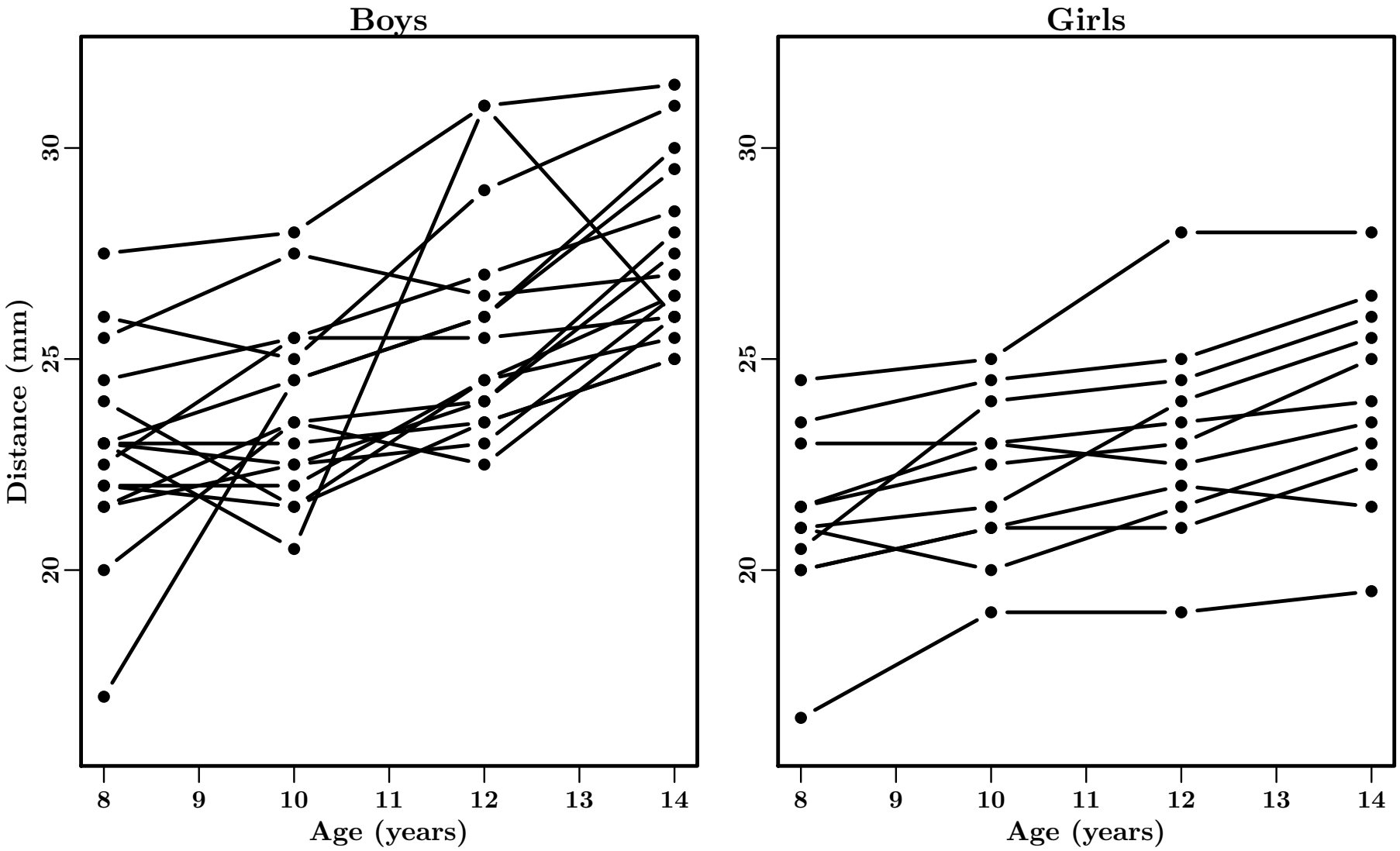
Group	Age 8	Age 10	Age 12	Age 14
Boys	26.0	25.0	29.0	31.0
	21.5	22.5	23.0	26.5
	⋮	⋮	⋮	⋮
Girls	21.0	20.0	21.5	23.0
	21.0	21.5	24.0	25.5
	⋮	⋮	⋮	⋮

Question: Are the growth profiles the same for boys and girls in the age interval considered in this investigation.

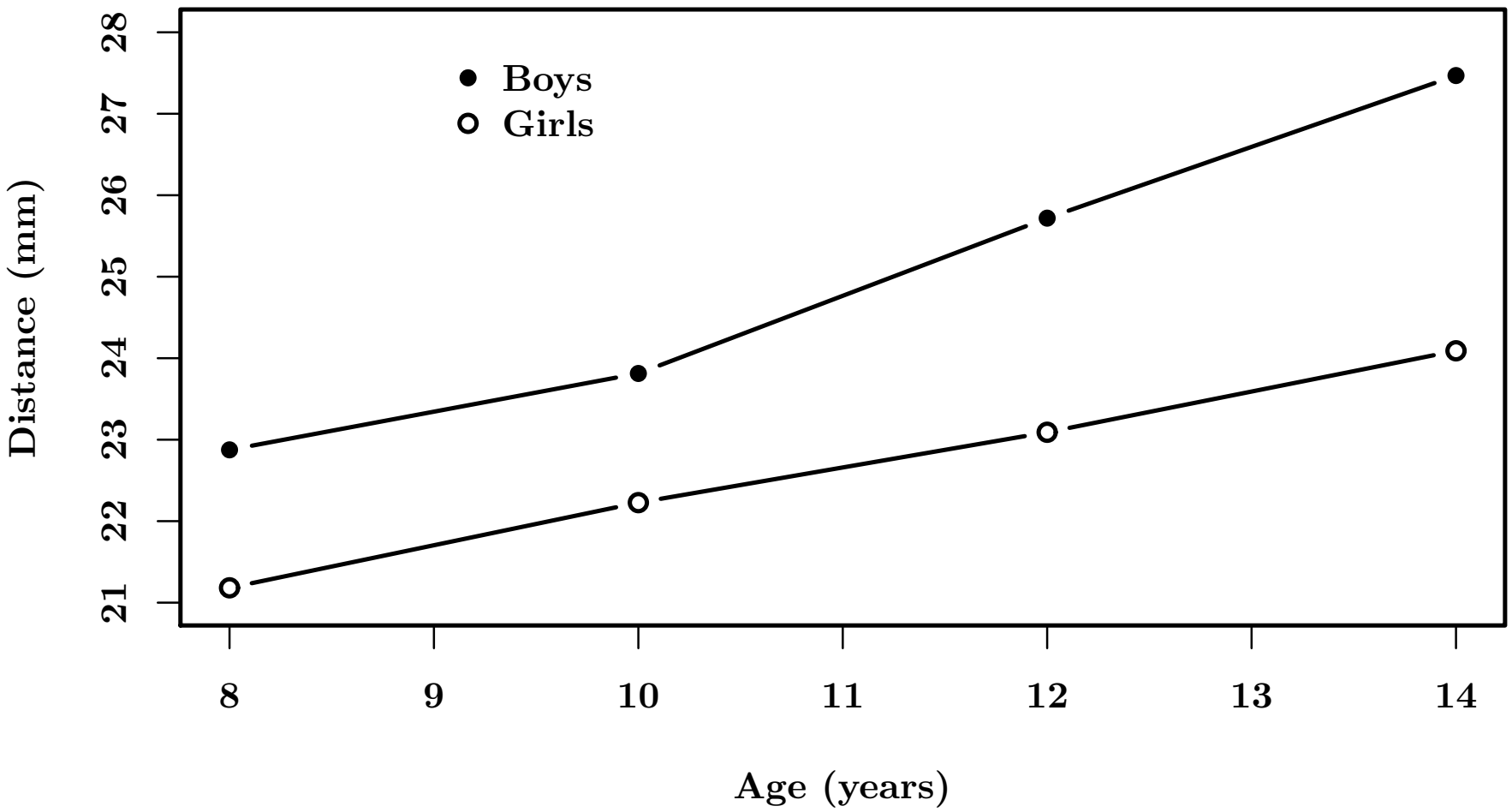
Orthodontic measurements: Individual curves



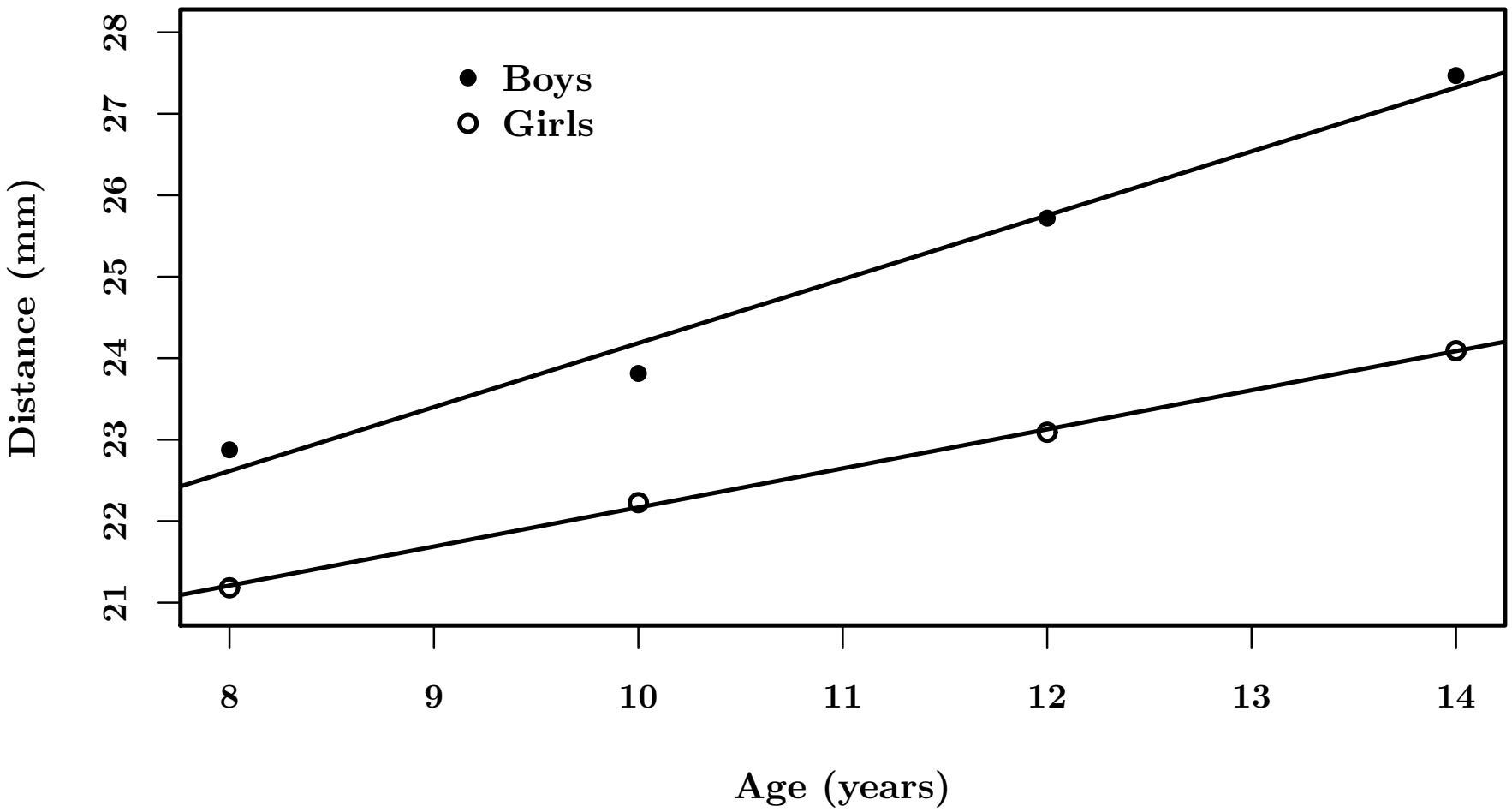
Orthodontic measurements: Grouped curves



Orthodontic measurements: Mean curves



Orthodontic measurements: Means and lines



Orthodontic measurements: Parameter estimates from the Random Coefficient Model

The estimated slopes and intercepts for the two populations:

$$\hat{\alpha}_{\mathtt{B}} = 24.97 \quad [23.93, 26.01], \qquad \hat{\beta}_{\mathtt{B}} = 0.78 \quad [0.59, 0.98]$$

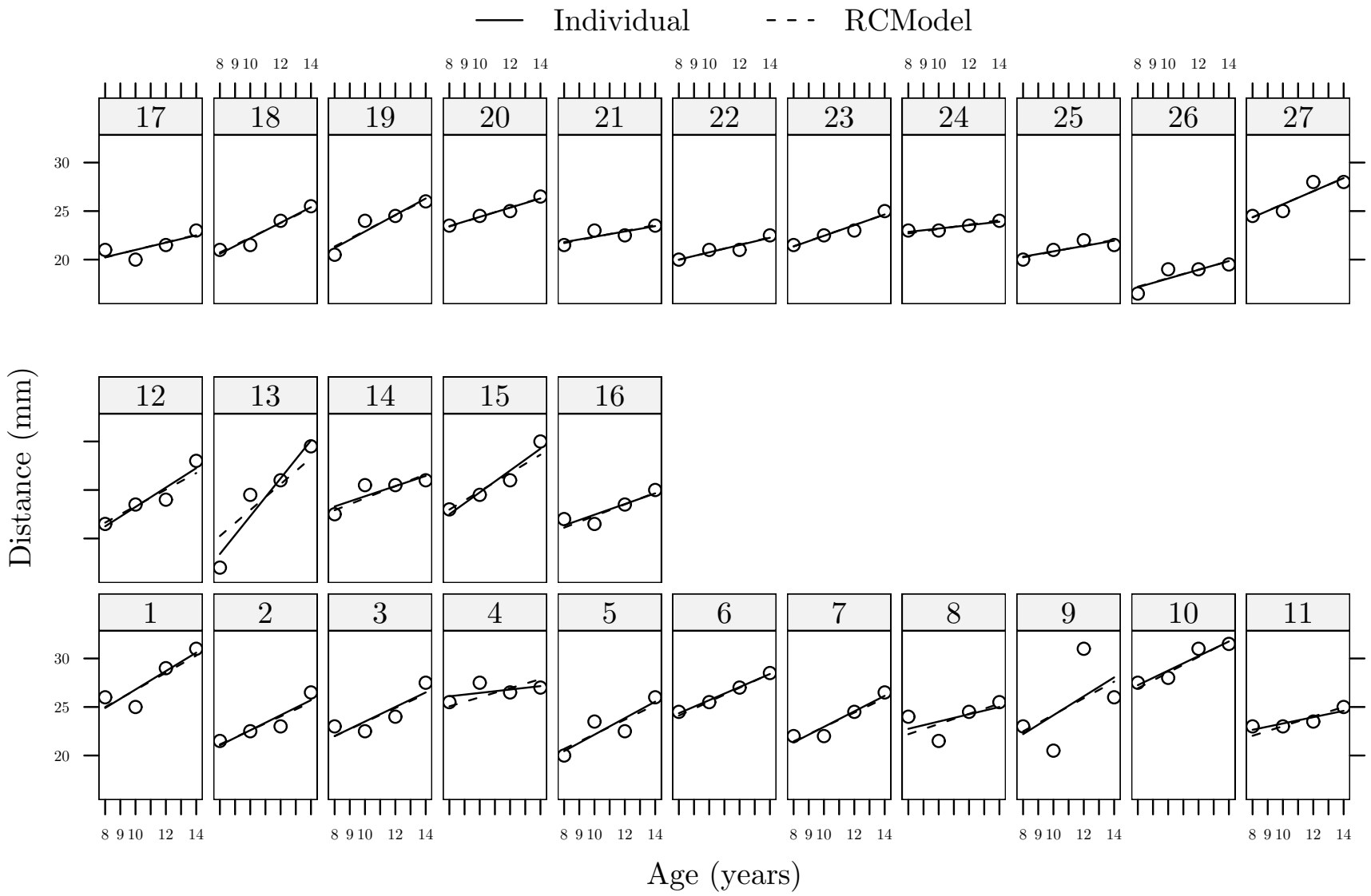
$$\hat{\alpha}_{\mathtt{G}} = 22.65 \quad [21.47, 23.83], \qquad \hat{\beta}_{\mathtt{G}} = 0.48 \quad [0.33, 0.63]$$

A **test for equal slopes** gives a p-value of 0.0154 and so we conclude that the **change over time is different for boys and girls**.

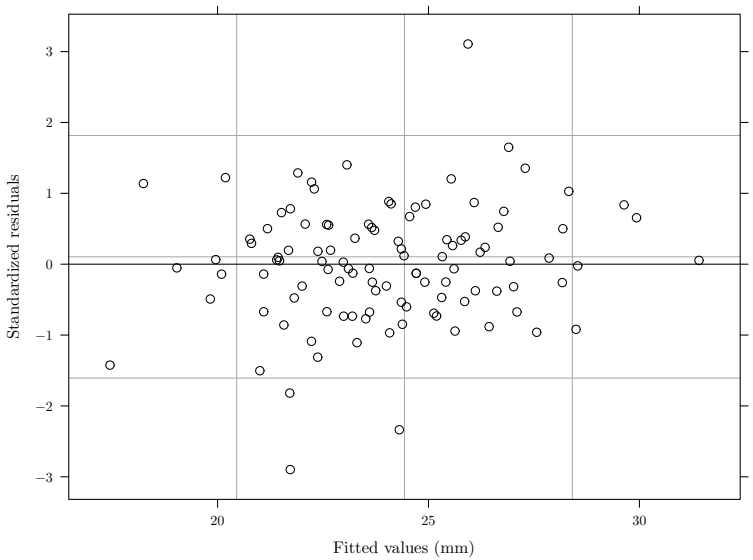
The standard deviation estimates are:

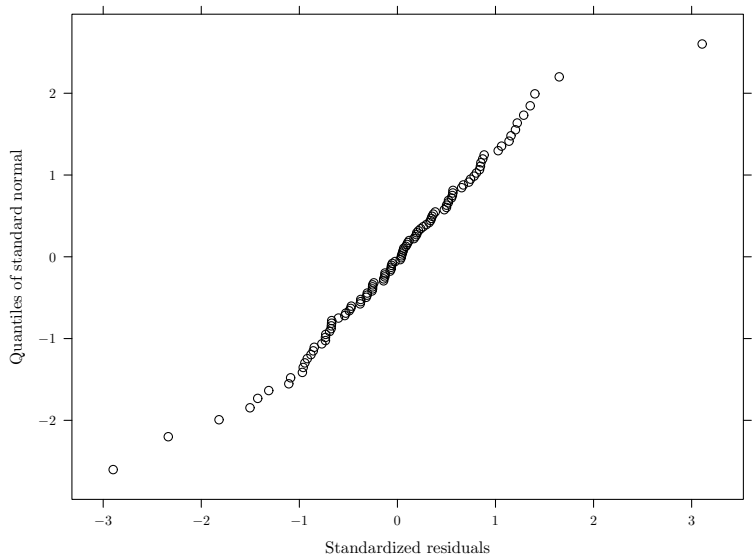
$$\hat{\sigma}_{\alpha} = 1.859, \qquad \hat{\sigma}_{\beta} = 0.165, \qquad \hat{\sigma}_{\mathtt{B}} = 1.613, \qquad \hat{\sigma}_{\mathtt{G}} = 0.668,$$

Orthodontic measurements for boys and girls: Model fit



Orthodontic measurements: Residuals and conclusions



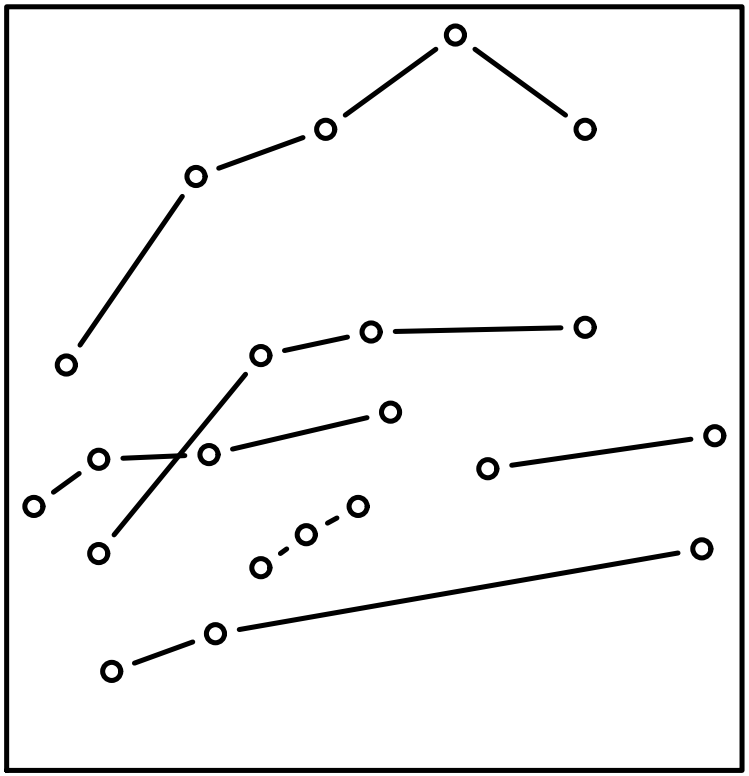


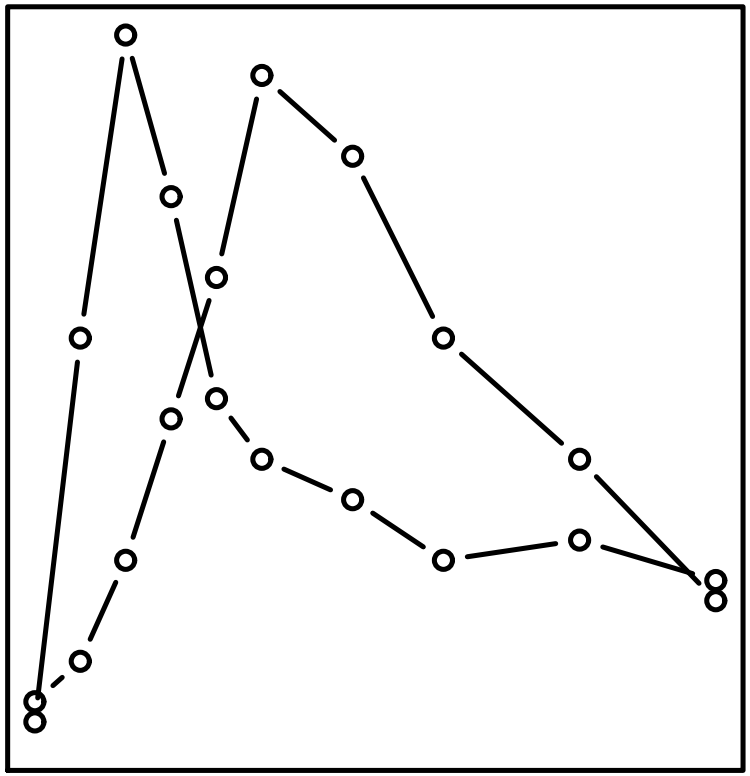
The estimated expected difference in distance between a boy and a girl both aged 11 years is

$$2.32 \text{ mm}, \qquad 95\% - \text{CI}: \ 0.75 \text{ mm to } 3.89 \text{ mm}$$

whereas for children aged 14 the estimated expected distance difference is 3.24 mm (95%-CI: 1.53 mm to 4.94 mm).

More general problems with repeated measurements I





More general problems with repeated measurements II

Two common situations where the MANOVA and univariate repeated measurements ANOVA analyzes cannot be used:

- **Irregularly spaced measurements**: For example visits to a GP
- **Horizontally shifted curves**: For example due to delayed reaction

One possible solution is to include the covariate directly in a repeated measurements regression analysis, either

- In a **linear** regression
- In a **non-linear** regression analysis: Maybe based on a **model** of the dynamical processes

Another is to consider a **summary statistic** (sum up each curve in a single quantity). More on this after lunch.

Repeated measurements in more complex designs

So far we have mainly focused on **comparing two or three groups** based on repeated measurements for each individual.

It should come as no surprise that the analysis of repeated measurements can be extended to more complex designs, such as:

- Designs with **more than one treatment** factor, for example two-way ANOVA where each measurement is a curve (repeated measurements).
- **Cross-over trials** where all patients are given both treatments and where the response is a curve.
- **Double repeated measurements** where patients are examined several times and each time the measurement is a repeated measurements curve.

Let us consider an example of a cross-over trial.

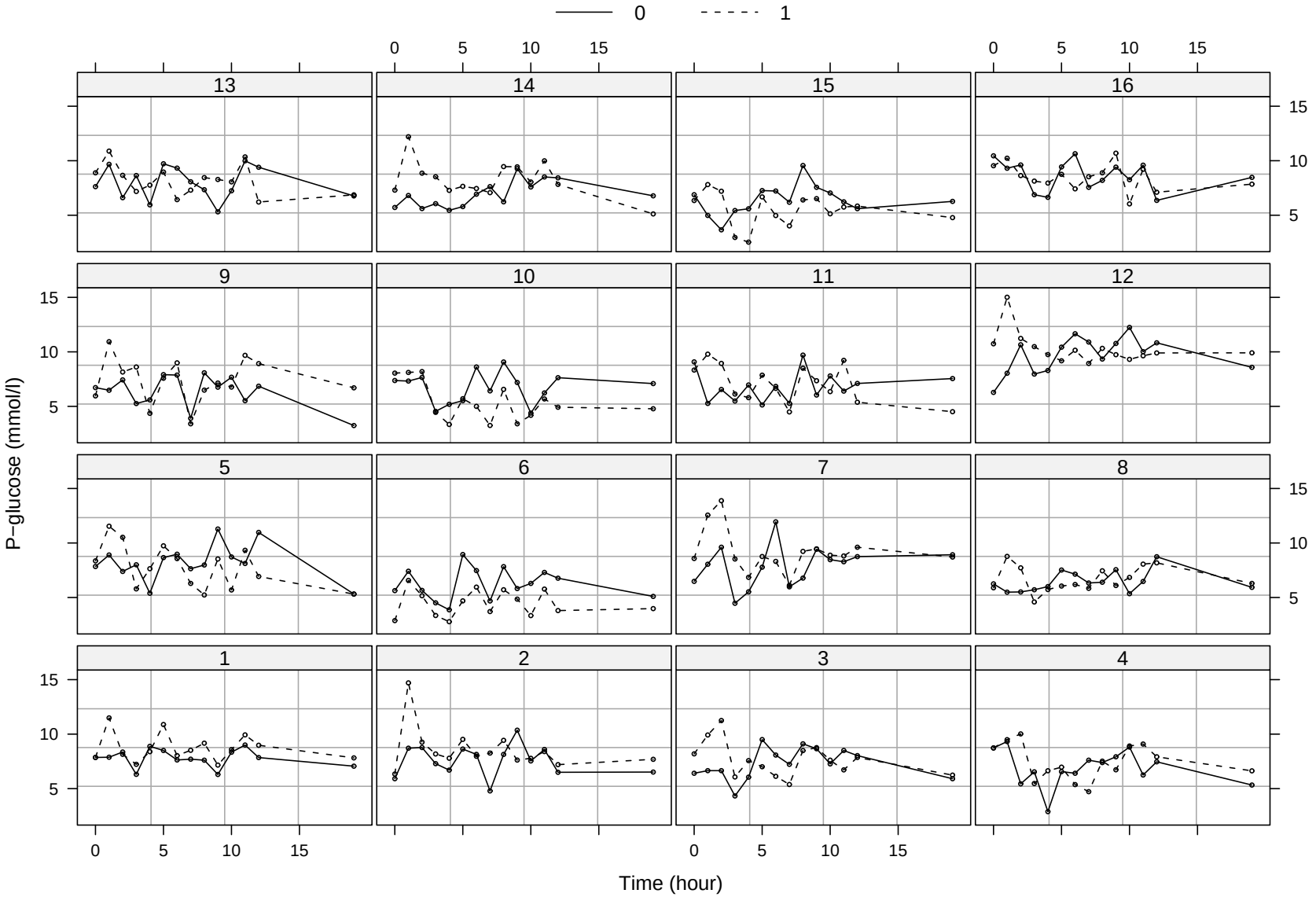
Example: Diet and plasma glucose for diabetic patients

Data: 16 patients with type 2 diabetes had their plasma glucose levels (mmol/l) measured 14 times during each of two consecutive days. The patients were randomized to either receiving a normal diet on day one followed by a diet with no breakfast on day two or the two diets in the reverse order.

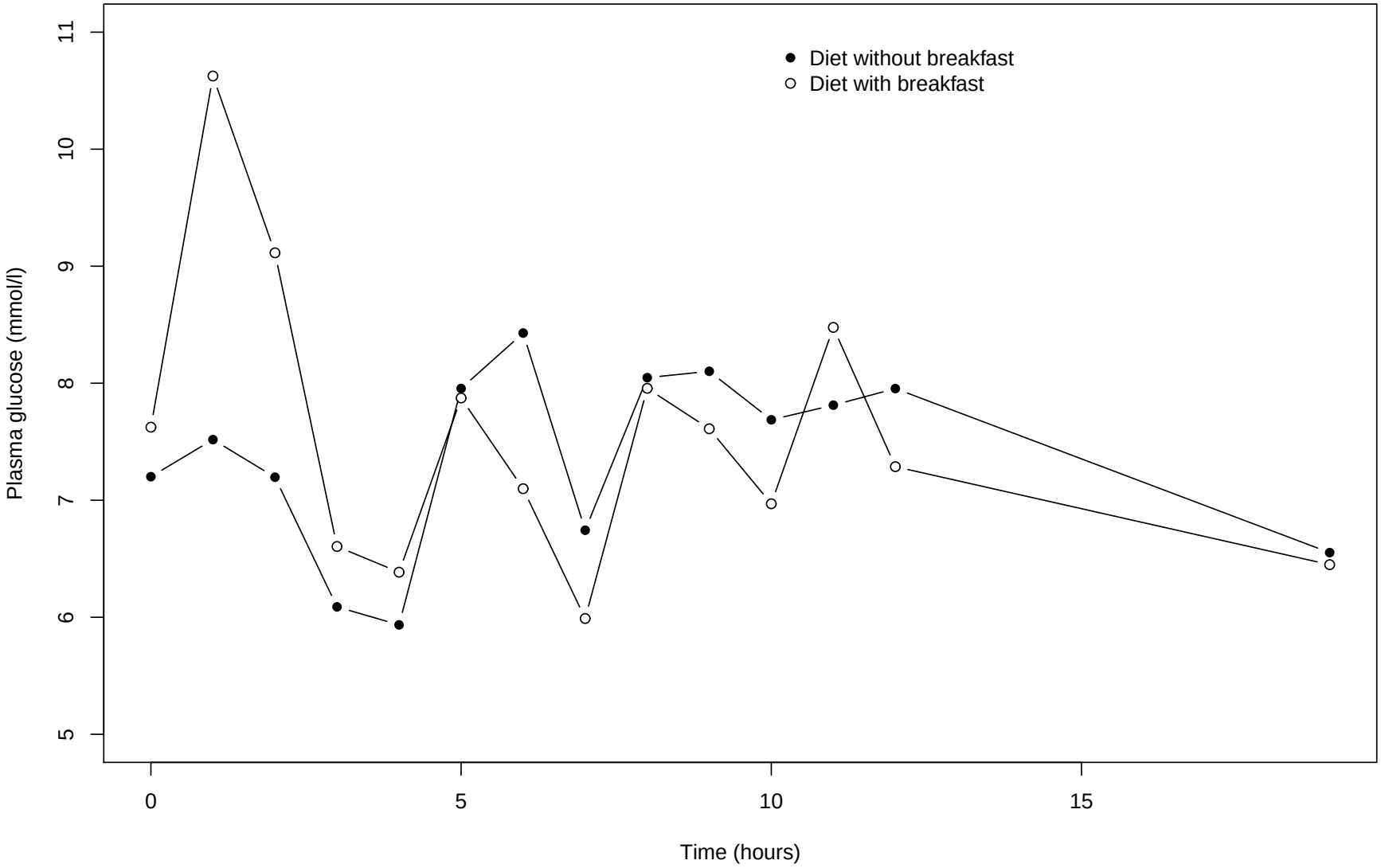
Patient	Day	Breakfast	8 am	9 am	10 am	$\cdots$	8 pm	3 am
1	1	0	7.86	7.88	8.36	$\cdots$	7.85	7.06
1	2	1	7.86	11.49	8.15	$\cdots$	8.98	7.82
2	1	0	5.91	8.71	8.77	$\cdots$	6.50	6.52
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$

Question: Is there an effect of the diet without breakfast on the plasma glucose level? Is the plasma glucose level more erratic for patients who receive breakfast?

Diet and glucose for diabetic patients: Individual curves



Diet and glucose for diabetic patients: Mean curves



Diet and glucose for diabetic patients: Design variables

Again we are interested in the usual factors:

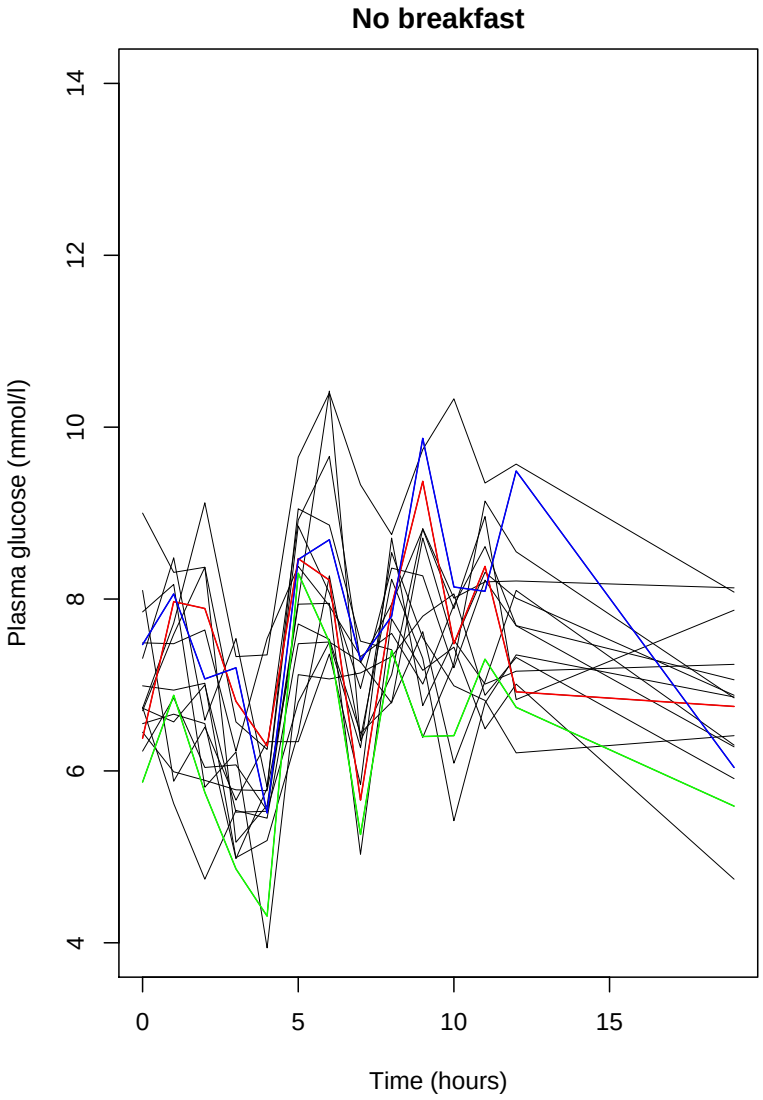
- **Treatment**: Breakfast / No breakfast
- **Time**: Hours since 8:00 am

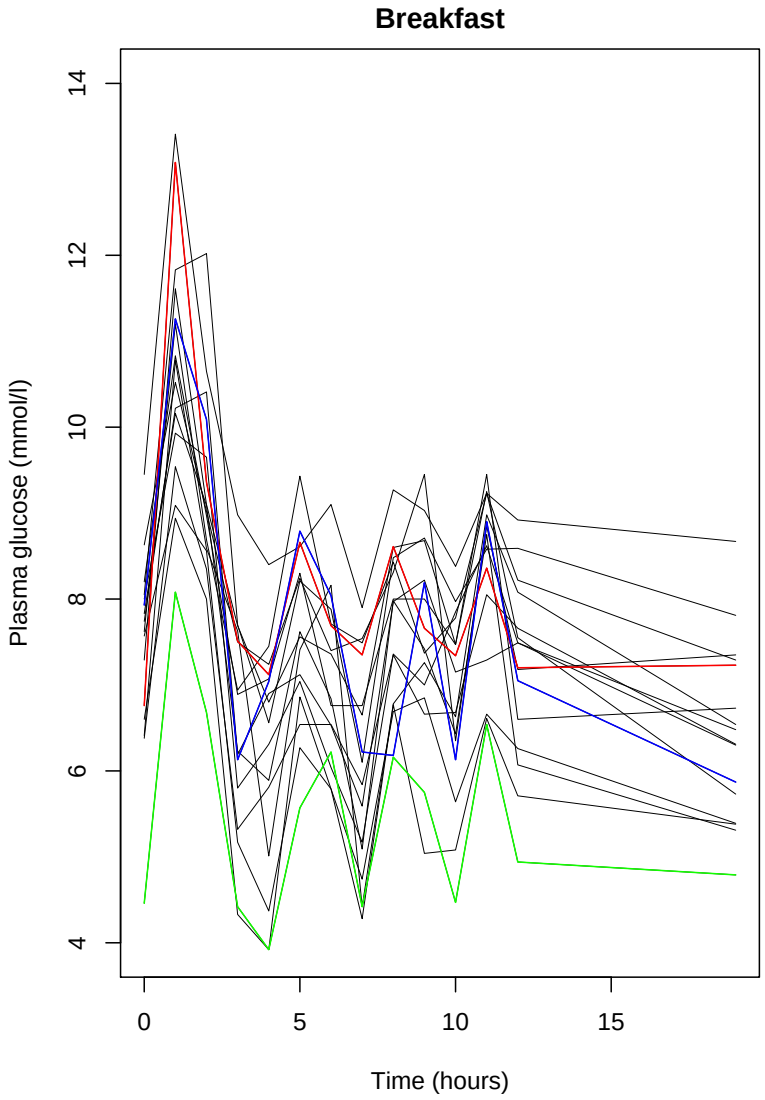
However, the nature of the design (cross-over trial) has to be taken into account. This means that we have to consider the factors:

- **Patient**: $1, 2, \ldots, 16$
- **Day**: 1 / 2
- **Order**: Breakfast - No breakfast / No breakfast - breakfast
- **Carry-over**: Previous treatment - 0, breakfast, no breakfast

Furthermore, we may wish to include some interactions in the analysis: For example, is the difference between patients the same in the treatment groups (if not we would include an interaction between patient and treatment).

Diet and plasma glucose: Individual curves for each diet





Diet and glucose: Analysis based on summary statistics

When dealing with more complex designs (and often also when dealing with simple designs) it can be very instructive to make an analysis based on a

- **Summary statistic:** A quantity calculated from each curve

A summary statistic has to reflect important aspects of the problem at hand:

- The **average** plasma glucose level
- **Area under the curve** (AUC)
- **Standard deviation** of the plasma glucose
- **MAGE** (mean amplitude of glycemic excursions)

Other popular choices of summary statistics include:

- **Increase** over a certain period
- The **maximum** value
- The **slope** in a certain period
- **Time until** the maximum value is attained

Diet and glucose: The mean and standard deviation

As examples of summary statistics let us consider the plasma glucose mean and standard deviation.

Mean plasma glucose

	Day 1	Day 2		Day 1	Day 2
	Breakfast			Breakfast	
Patient	No	Yes	Patient	Yes	No
1	7.57	8.20	9	7.33	6.63
2	7.46	8.14	10	6.12	6.73
3	7.17	7.51	11	7.19	6.89
4	6.90	7.34	12	9.29	8.82
5	7.80	7.70	13	7.86	7.62
6	6.33	5.46	14	7.95	7.03
7	7.62	8.46	15	6.19	6.56
8	6.68	6.93	16	8.11	8.05

Diet and plasma glucose: Analysis of the mean

In the analysis of the mean plasma glucose, we are interested in whether it depends on the **treatment** (breakfast or no breakfast), but we want to take into account:

- **Patient**: Patients have different natural plasma glucose levels
- **Order**: The response may depend on the order in which the patient receives the two treatments
- **Day**: There may be some systematic variation between days in a hospital

In this case we get a mean plasma glucose level that is

$$0.25 \text{ mmol/l}, \quad 95\% - \text{CI}: \ -0.04 - 0.53 \text{ mmol/l}$$

higher having had breakfast compared to not having had breakfast.

The test for no effect of the treatment results in a p-value of 0.08, and we conclude that there is no significant effect of having no breakfast on the mean plasma glucose level.

Diet and plasma glucose: The mean before and after lunch

Even though there is no significant overall difference in the mean plasma glucose levels, there is a big difference before lunch:

Before lunch: 1.42 mmol/l, 95%−CI: 1.03 − 1.82 mmol/l, $p < 0.0001$
After lunch: −0.41 mmol/l, 95%−CI: −0.71 − −0.11 mmol/l, $p = 0.0108$

Conclusion:

- The overall mean plasma glucose is not influenced significantly by not having breakfast (breakfast - no breakfast: 0.25 mmol/l, 95%−CI: $-0.04 - 0.53$ mmol/l, $p = 0.08$).
- Before lunch the mean plasma glucose is significantly higher after breakfast compared to no breakfast (1.42 mmol/l, 95%−CI: $1.03 - 1.82$ mmol/l, $p < 0.0001$).
- After lunch the mean plasma glucose is significantly higher after no breakfast compared to after breakfast (0.41 mmol/l, 95%−CI: $0.11 - 0.71$ mmol/l, $p = 0.0108$).

Diet and plasma glucose: The standard deviation

The degree of erratic behaviour is to some extend captured by the standard deviation.

Standard deviation of the plasma glucose

	Day 1	Day 2		Day 1	Day 2
	Breakfast			Breakfast	
Patient	No	Yes	Patient	Yes	No
1	0.59	1.14	9	1.61	1.06
2	1.03	1.60	10	1.47	1.00
3	1.12	1.39	11	1.48	0.88
4	1.13	1.39	12	1.35	1.18
5	1.18	1.64	13	1.24	1.05
6	1.12	1.17	14	1.43	0.91
7	1.35	1.64	15	1.36	1.02
8	0.79	1.14	16	1.10	1.02

Diet and plasma glucose: The SD before and after lunch

Again we see big differences in SD before and after lunch (B - no B):

Overall:	0.36 mmol/l,	95%−CI:	$0.26 - 0.46$ mmol/l, $p < 0.0001$
Before lunch:	1.07 mmol/l,	95%−CI:	$0.82 - 1.32$ mmol/l, $p < 0.0001$
After lunch:	0.01 mmol/l,	95%−CI:	$-0.13 - 0.16$ mmol/l, $p = 0.83$

Conclusion:

- The overall standard deviation is significantly larger after breakfast compared to no breakfast (0.36 mmol/l, 95%−CI: $0.26 - 0.46$ mmol/l, $p < 0.0001$).
- Before lunch the standard deviation is significantly higher after breakfast compared to no breakfast (1.07 mmol/l, 95%−CI: $0.82 - 1.32$ mmol/l, $p < 0.0001$).
- After lunch the standard deviation is not significantly influenced by having breakfast (0.01 mmol/l, 95%−CI: $-0.13 - 0.16$ mmol/l, $p = 0.83$).

Diet and plasma glucose: The analyzes in Stata

In **Stata** the analyzes of summary statistics can be done in the following way:

```
egen mpgluc = mean(pgluc), by(pt day)
reshape wide pgluc, i(pt day) j(time)

anova mpgluc breakfast pt day order
pwcompare breakfast, eff
                          |Contrast Std. Err.  t  P>|t| [95% Conf.Interval]
--------------------------+-----------------------------------------------
breakfast                 |
Breakfast vs No breakfast|.2456696 .1324916 1.85 0.085 -.0384966  .5298358
```

Note: Look only at the test corresponding to `breakfast` as for example the effect of `order` has to be tested against the variation between patients.

```
anova mpgluc order / pt breakfast day
```

Choice of summary statistics

There are a number of considerations that you have to take into account when choosing a summary statistic:

- The summary statistic **has to reflect important aspects** of the medical problem at hand
- The summary statistic should be chosen **BEFORE** you look at the data (should be written in the protocol)
- You can easily consider **several summary statistics**
- You **throw away a lot of information**, when replacing a whole curve with a summary statistic, so one should consider if the primary scientific question could be answered based on an analysis of all the data
- Factors that characterize the **experimental design** and the **experimental circumstances** still have to be included in the analyses

Diet and plasma glucose: Analysis of all the data

Factors of interest:

- **Treatment**: Breakfast / No breakfast
- **Time**: Hours since 8:00 am

Design variables, not of primary interest but should be taken into account:

- **Patient**: $1, 2, \ldots, 16$
- **Day**: 1 / 2
- **Order**: Breakfast - No breakfast / No breakfast - breakfast

Since we have 14 time points, it is not really feasible to analyze the data using a **multivariate repeated measurement analysis** (this would require estimating 105 standard deviations and correlations).

The **univariate repeated measurement analysis** only requires the estimation of one standard deviation and one correlation. There are models in between the two, as we saw in the last example on day 2.

Diet and glucose: Results from the analysis of all the data

The test for parallel mean curves for the two diets gives a p-value of

$$p < 0.0001$$

Remarks:

- The different corrections to the p-value does not alter it to any great extend
- A test for reducing the **multivariate repeated measurement model** to the **univariate repeated measurement model** gives:

  Breakfast: $p = 0.47$
  No breakfast: $p = 0.80$

  But a reduction from 105 to 2 parameters results in a **VERY** weak test.

Conclusion: There is clear evidence in the data against the hypothesis of equal development in the plasma glucose after breakfast and no breakfast (no surprise).

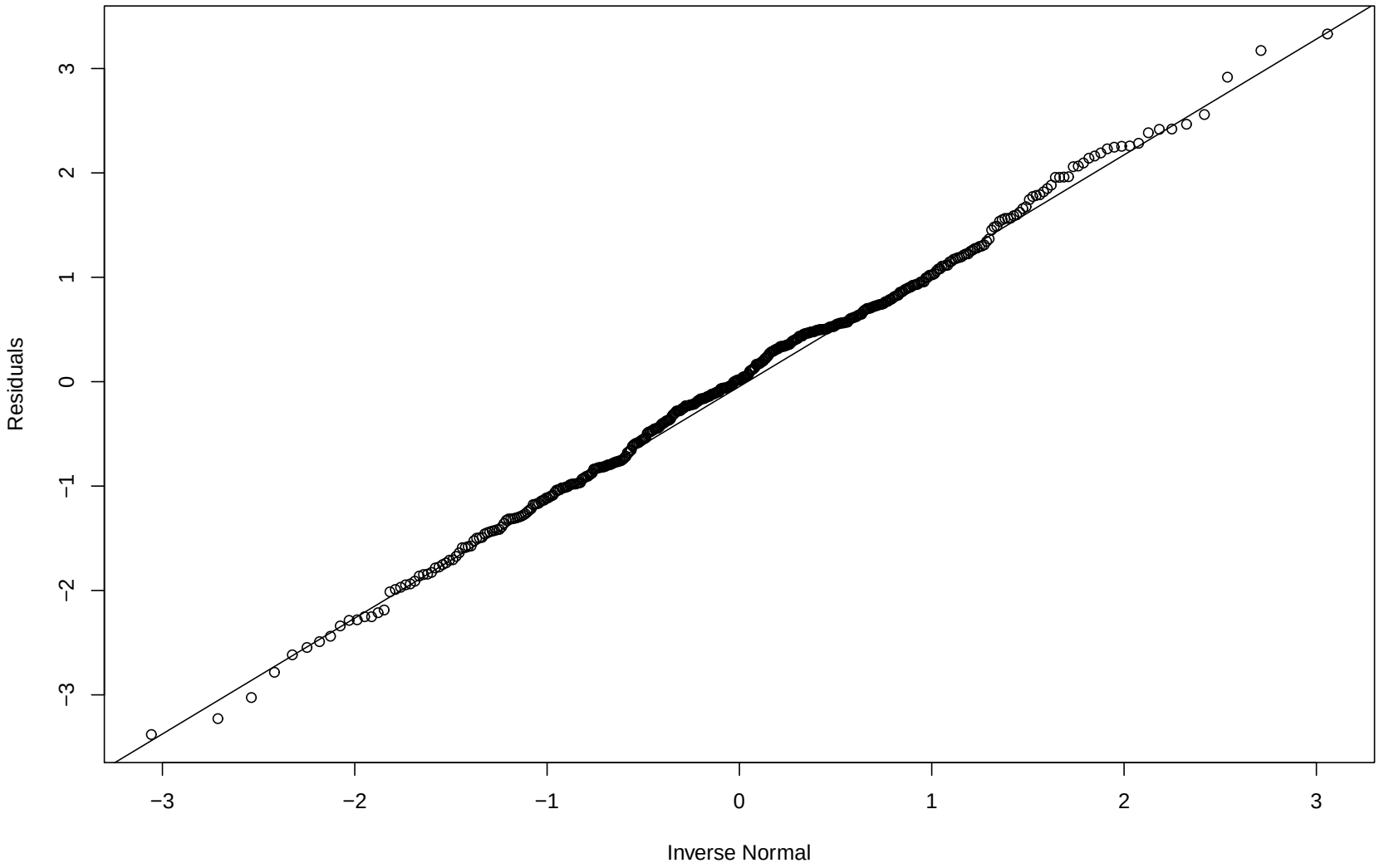
Diet and glucose: Analysis of all the data in Stata

In **Stata** the analysis of all the data can be done in the following way:

```
anova pgluc order /pt day breakfast /pt|day ///
            time breakfast#time, bse(pt#day) repeated(time)
```

Source	Partial SS	df	MS	F	Prob > F
order	.532350565	1	.532350565	0.03	0.8675
pt	258.045255	14	18.4318039		
day	.108439354	1	.108439354	0.06	0.8177
breakfast	1.74311433	1	1.74311433	0.89	0.3624
pt|day	27.5246996	14	1.96604997		
time	283.663079	13	21.8202369	46.01	0.0000
breakfast#time	162.269941	13	12.4823032	26.32	0.0000
Residual	184.95717	390	.474249155		

Diet and plasma glucose for diabetic patients: Residuals



Analysis of more complicated repeated measurements designs: Recommendations I

A few recommendations when contemplating setting up a more complicated design that involves repeated measurements:

Before you have the data:

- Formulate the questions that you want to answer based on the statistical analysis of the data.
- Target the experiment to enable you to answer the questions of interest.
- Write down the strategy for analyzing the data - this could involve summary statistics and/or more elaborate analyzes (in fact the **Stata** program could to a large extend be written in advance).
- Make sure that you make a note of any special circumstances that occur during the experiment.

Analysis of more complicated repeated measurements designs: Recommendations II

Before you have the data:

- If the **overall change** is of interest then place observation time-points at the start and late (to ensure that steady state has been reached).
- If **how you got to the final level** is of interest then place the observation time-points close when things are expected to happen.

After you have the data:

- **Follow the analysis plan!**
- If you have a **model** for the evolvement over time that should be included in the analysis.
- Take into account **unforeseen factors** that could influence the results.
- **Get help** if you are unsure whether the analysis is correct (these are very complicated analyzes but only because the data require them to be).